

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
**Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 13:37:13 ; Search time 3200 Seconds

(without alignments)  
10558.862 Million cell updates/sec

Title: US-10-054-562a-3

Perfect score: 1161

Sequence: 1 atgatgatctctctatctgc.....ttcgtccaatcagaagca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_pln:\*  
35: em\_hlg\_rod:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vrt:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1161	100.0	1161	6	AR204520	Sequence
2	1161	100.0	1161	6	AR204521	Sequence
3	1161	100.0	1779	6	AR204518	Sequence
4	1161	100.0	1779	6	AR204519	Sequence
5	646.4	55.7	1380	3	AF125580	Wuchereri
6	504.2	43.4	1372	6	AR204522	Sequence
7	504.2	43.4	1372	6	AR204523	Sequence
8	461	39.7	892	6	AR204531	Sequence
9	461	39.7	892	6	AR204532	Sequence
10	339.2	29.2	25823	3	CEP2285	Sequence
11	323.8	27.9	34831	3	CEC47G2	Sequence
12	315	27.1	2310	3	CELCUT1	Sequence
13	310.4	26.7	813	6	AR204524	Sequence
14	310.4	26.7	813	6	AR204525	Sequence
15	249	21.4	5792	3	ALU73005	Sequence
16	217.2	18.7	2584	3	MAMTCUT1	Sequence
17	190.6	16.4	637	3	BMA012617	Sequence
18	156.6	13.5	358	3	BP0012618	Sequence
19	107.4	9.3	39478	3	CEFS3F1	Sequence
20	85.6	7.4	56511	3	AC084467	Caenorhabd
21	83.8	7.2	31536	3	CEFB04D5	Caenorhabd
22	74.8	6.4	32412	3	CEFB3B6	Caenorhabd
23	67	5.8	39736	3	CEB07E3	Caenorhabd
24	62.6	5.4	36306	3	CEM142	Caenorhabd
25	62.2	5.4	36089	3	CEB265	Caenorhabd
26	56.2	4.8	25284	3	CEW01A8	Caenorhabd
27	55.4	4.8	59533	3	AC084450	Caenorhabd
28	54.8	4.7	28396	3	U41264	Caenorhabd
29	53.6	4.6	40265	3	CEM28	Caenorhabd
30	52.8	4.5	34368	3	CEM06D12	Caenorhabd
31	51.8	4.5	40600	3	CEB20D1	Caenorhabd
32	49	4.2	26449	3	U23449	Caenorhabd
33	48.4	4.2	31731	3	U21318	Caenorhabd
34	48.4	4.2	327650	1	MPULM01	Caenorhabd
35	48.2	4.2	53150	2	AC116984	Caenorhabd
36	47	4.0	45713	3	AF067608	Dictyoste
37	46.8	4.0	27770	3	CEB22C8	Caenorhabd
38	44.8	3.9	6942	1	AF455359	Sp1roplas
39	44	3.8	229420	2	AL772373	Danio rer
40	43	3.7	81091	9	AL645495	Human DNA
41	43	3.7	164389	3	PFMAL3P6	Caenorhabd
42	43	3.7	165318	2	AC126747	Rattus no
43	42.8	3.7	48063	8	AF288090	Rhodomon
44	42.2	3.6	1141	6	AX083744	Sequence
45	41.4	3.6	79769	2	AC115678	Dictyoste

## ALIGNMENTS

RESULT 1  
AR204520  
LOCUS AR204520 1161 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 3 from patent US 6368600.  
ACCESSION AR204520  
VERSION AR204520.1 GI:21501867  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1161)  
AUTHORS Chandrasekhar, R. and Morales, T. H.  
TITLE Parasitic helminth cuticlin nucleic acid molecules and uses thereof  
JOURNAL Patent: US 6368600-A 3 09-APR-2002;  
FEATURES Location/Qualifiers

source 1.1161  
/organism="unknown"  
BASE COUNT 338 a 227 c 235 g 361 t  
ORIGIN

Query Match 100.0%; Score 1161; DB 6; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 6.3e-262;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATGATTCGCTTATGCTTCTGCTACCTATTCGATTCGCTTATTCGATTCGCG 60
   |||
DB 1 ATGATATTCGCTTATGCTTCTGCTACCTATTCGATTCGCTTATTCGATTCGCG 60
OY 61 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGTAATGTGAGCAACTTCAATACATC 120
   |||
DB 61 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGTAATGTGAGCAACTTCAATACATC 120
OY 121 AATTTTAATACAGTATGATTCGATTCGAGACATGTTTATGTGAAGGCTTTATGATCAA 180
   |||
DB 121 AATTTTAATACAGTATGATTCGATTCGAGACATGTTTATGTGAAGGCTTTATGATCAA 180
OY 181 GAAGGTGGCGTAATGATGATGAGGTGAGCGTCAAGTTGCCGGAATTCCTTCATTTGAT 240
   |||
DB 181 GAAGGTGGCGTAATGATGATGAGGTGAGCGTCAAGTTGCCGGAATTCCTTCATTTGAT 240
OY 241 TCATGCAATGTTGCCGTAACAGATCTCTGAATCCACGTGTATTTTGTGAACAACACT 300
   |||
DB 241 TCATGCAATGTTGCCGTAACAGATCTCTGAATCCACGTGTATTTTGTGAACAACACT 300
OY 301 GTTGTATATTCGTTTCATCCATTTATTTGTACCAAGTTGATGTCGATATGATGATCAA 360
   |||
DB 301 GTTGTATATTCGTTTCATCCATTTATTTGTACCAAGTTGATGTCGATATGATGATCAA 360
OY 361 TCGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
   |||
DB 361 TCGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
OY 421 ACTGCTTTTCAAACTCAAAATGTTCCGATGCCAGTATGCCGTTATGAATTTTGTGATGCT 480
   |||
DB 421 ACTGCTTTTCAAACTCAAAATGTTCCGATGCCAGTATGCCGTTATGAATTTTGTGATGCT 480
OY 481 GGACCAACCGGTGCAACCACTCAATTTGCTATCATTGCTGACGACGATTTATCAATAATG 540
   |||
DB 481 GGACCAACCGGTGCAACCACTCAATTTGCTATCATTGCTGACGACGATTTATCAATAATG 540
OY 541 ACATGGATTCGAAACCGTGTATCTTCTGCGGTTGCTCATCTCCGTTTGTGCGAT 600
   |||
DB 541 ACATGGATTCGAAACCGTGTATCTTCTGCGGTTGCTCATCTCCGTTTGTGCGAT 600
OY 601 GATGGTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
   |||
DB 601 GATGGTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
OY 661 TTGCTTAATATTTTGGATATTCACACAGATTTAATGCTGCGCAAGAGCTCAGATATC 720
   |||
DB 661 TTGCTTAATATTTTGGATATTCACACAGATTTAATGCTGCGCAAGAGCTCAGATATC 720
OY 721 AATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
   |||
DB 721 AATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
OY 781 CCAAAATAGCAATGTTTGCACACAAATGTTGAGAACCAAGATTCGAGAGCTGTAA 840
   |||
DB 781 CCAAAATAGCAATGTTTGCACACAAATGTTGAGAACCAAGATTCGAGAGCTGTAA 840
OY 841 ACAGGTGGTGGCGGACCAAAACCTGCTGAGCTGCGCAACTTCGTTTACTCAAGAAAGA 900
   |||
DB 841 ACAGGTGGTGGCGGACCAAAACCTGCTGAGCTGCGCAACTTCGTTTACTCAAGAAAGA 900
OY 901 TCTGAGAACCGGAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 960
   |||
DB 901 TCTGAGAACCGGAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 960

```

OY 961 GATGATATATCAAGCTTTGGCAGTGTGATTTACGTACCGGTGACCTTCTGCAACATAATGA 1020  
|||  
DB 961 GATGATATATCAAGCTTTGGCAGTGTGATTTACGTACCGGTGACCTTCTGCAACATAATGA 1020

OY 1021 CAACCTGTAATACCTTCTGCTGACGATGACAAATGGAATGTGATGTCACCATTTGGCTTCTCA 1080  
|||  
DB 1021 CAACCTGTAATACCTTCTGCTGACGATGACAAATGGAATGTGATGTCACCATTTGGCTTCTCA 1080

OY 1081 ATGTTATGAGGTTTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
|||  
DB 1081 ATGTTATGAGGTTTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

OY 1141 TTTGCTCAAAATCAGAAAGCA 1161  
|||  
DB 1141 TTTGCTCAAAATCAGAAAGCA 1161

RESULT 2  
AR204521/c AR204521 1161 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 5 from patent US 6368600.  
ACCESSION AR204521  
VERSION AR204521.1 GI:21501869  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 1161)  
AUTHORS ChandraShekar, R. and Morales, T. H.  
TITLE Parasitic helminth cuticlin nucleic acid molecules and uses thereof  
JOURNAL Patent: US 6368600-A 5 09-APR-2002;  
FEATURES Location/Qualifiers  
source 1.1161  
/organism="unknown"

BASE COUNT 361 a 235 c 227 g 338 t  
ORIGIN

Query Match 100.0%; Score 1161; DB 6; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 6.3e-262;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATGATTCGCTTATGCTTCTGCTACCTATTCGATTCGCTTATTCGATTCGCG 60
   |||
DB 1161 ATGATGATTCGCTTATGCTTCTGCTACCTATTCGATTCGCTTATTCGATTCGCG 1102
OY 61 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGTAATGTGAGCAACTTCAATACATC 120
   |||
DB 1101 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGTAATGTGAGCAACTTCAATACATC 1042
OY 121 AATTTTAATACAGTATGATTCGATTCGAGACATGTTTATGTGAAGGCTTTATGATCAA 180
   |||
DB 1041 AATTTTAATACAGTATGATTCGATTCGAGACATGTTTATGTGAAGGCTTTATGATCAA 982
OY 181 GAAGGTGGCGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
   |||
DB 981 GAAGGTGGCGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 922
OY 241 TCATGCAATGTTGCCGTAACAGATCTCTGAATCCACGAGTGTATTTGTGAACAACACT 300
   |||
DB 921 TCATGCAATGTTGCCGTAACAGATCTCTGAATCCACGAGTGTATTTGTGAACAACACT 862
OY 301 GTTGTATATTCGTTTCATCCATTTATTTGTACCAAGTTGATGATGATGATGATGATGAT 360
   |||
DB 861 GTTGTATATTCGTTTCATCCATTTATTTGTACCAAGTTGATGATGATGATGATGATGAT 802
OY 361 TCGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
   |||
DB 801 TCGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 742
OY 421 ACTGCTTTTCAAACTCAAAATGTTCCGATGCCAGTATGCCGTTATGAATTTTGTGATGCT 480
   |||
DB 741 ACTGCTTTTCAAACTCAAAATGTTCCGATGCCAGTATGCCGTTATGAATTTTGTGATGCT 682

```

QY 481 GGACCAACCGGTCACACAGTTCATTTGCTATCATTTGGTCAGCCAGTTTATCATTAATGG 540  
|||||  
Db 681 GGACCAACCGGTCACACAGTTCATTTGCTATCATTTGGTCAGCCAGTTTATCATTAATGG 622  
QY 541 ACATGCGATTCGTAACCGTTGATCTTCTGCGCGTTGCTCATTCCTGCTTGTGAT 600  
|||||  
Db 621 ACATGCGATTCGTAACCGTTGATCTTCTGCGCGTTGCTCATTCCTGCTTGTGAT 562  
QY 601 GATGTAACGGTGTACTGTGGAAATCTTAAATGCTATGATGATGCTGCTTGAATAT 660  
|||||  
Db 561 GATGTAACGGTGTACTGTGGAAATCTTAAATGCTATGATGATGCTGCTTGAATAT 502  
QY 661 TTGCTAAATATTTGGAATATCCAAAGATTTAATGGTCGCCAAGAAGCTCAGTATAC 720  
|||||  
Db 501 TTGCTAAATATTTGGAATATCCAAAGATTTAATGGTCGCCAAGAAGCTCAGTATAC 442  
QY 721 AATATGCGGATGATCATCAGCTTTTCTATCATTCAGATCAGTATTTACCATTAAGAA 780  
|||||  
Db 441 AATATGCGGATGATCATCAGCTTTTCTATCATTCAGATCAGTATTTACCATTAAGAA 382  
QY 781 CCAATATGCGGATGATCATCAGCTTTTCTATCATTCAGATCAGTATTTACCATTAAGAA 840  
|||||  
Db 381 CCAATATGCGGATGATCATCAGCTTTTCTATCATTCAGATCAGTATTTACCATTAAGAA 322  
QY 841 ACAAGTGTGCGCGACGCAAAACCTGCTGACAGCTGCGCACTGCTTACTCAAGAAAGA 900  
|||||  
Db 321 ACAAGTGTGCGCGACGCAAAACCTGCTGACAGCTGCGCACTGCTTACTCAAGAAAGA 262  
QY 901 TCTGCAACACCGGAGATATCATTTGATGACGAACTGATATCAACACCTTGAATTAAGC 960  
|||||  
Db 261 TCTGCAACACCGGAGATATCATTTGATGACGAACTGATATCAACACCTTGAATTAAGC 202  
QY 961 GATGATATGAGGTTGGCAAGTTGATTTAGCTGACCGTCACTGCTGCAACATTAAGGA 1020  
|||||  
Db 201 GATGATATGAGGTTGGCAAGTTGATTTAGCTGACCGTCACTGCTGCAACATTAAGGA 142  
QY 1021 CAACCTGTAATCTTGTGTCAGTACAAATGATGATGATGATGATGATGATGATGATGAT 1080  
|||||  
Db 141 CAACCTGTAATCTTGTGTCAGTACAAATGATGATGATGATGATGATGATGATGATGAT 82  
QY 1081 ATGTTTATGCGTTTAAAGCATTCGATTTGCTGCGCTCATTTTAAACATTTGCTTAAA 1140  
|||||  
Db 81 ATGTTTATGCGTTTAAAGCATTCGATTTGCTGCGCTCATTTTAAACATTTGCTTAAA 22  
QY 1141 TTTCGTCGAATCAGAAGCA 1161  
|||||  
Db 21 TTTCGTCGAATCAGAAGCA 1

RESULT 3  
AR204518 1779 bp DNA linear PAT 20-JUN-2002  
LOCUS AR204518  
DEFINITION Sequence 1 from patent US 6368600.  
ACCESSION AR204518  
VERSION AR204518.1 GI:21501865  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1779)  
AUTHORS Chaudrasherkar, R. and Morales, T. H.  
TITLE Parasitic helminth cuticulin nucleic acid molecules and uses thereof  
JOURNAL Patent: US 6368600-A 1 09-APR-2002;  
FEATURES  
source 1.1779  
location/Qualifiers  
BASE COUNT 575 a 327 c 302 g 575 t  
ORIGIN

Query Match 100.0%; Score 1161; DB 6; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 6.3e-262;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGATGCGTCTTATTTGCTTCTGCTTACTACACTTATTCGATTTGCTTATTCGATTCG 60  
|||||  
Db 167 ATGATGATGCGTCTTATTTGCTTCTGCTTACTACACTTATTCGATTTGCTTATTCGATTCG 226  
QY 61 GTTGACAAATGCTGTGCAAGGTGAGCCAGAAATTTGAATGTGACCACTTAAATTAACATC 120  
|||||  
Db 227 GTTGACAAATGCTGTGCAAGGTGAGCCAGAAATTTGAATGTGACCACTTAAATTAACATC 286  
QY 121 AATTTTAAATACAGTAATTCATTCGAAAGACATGTTTATGTAAGAGTCTTTATGATCAA 180  
|||||  
Db 287 AATTTTAAATACAGTAATTCATTCGAAAGACATGTTTATGTAAGAGTCTTTATGATCAA 346  
QY 181 GAAGGTGCGGTATGATGAAAGGTGAGCGTCAAGTTGCCGAATTCCTTCCATTTGAT 240  
|||||  
Db 347 GAAGGTGCGGTATGATGAAAGGTGAGCGTCAAGTTGCCGAATTCCTTCCATTTGAT 406  
QY 241 TCATGCAATGTTGCGGTGATACAGATCTCTAATATCAGAGTGTATTTTGTGAACAACAT 300  
|||||  
Db 407 TCATGCAATGTTGCGGTGATACAGATCTCTAATATCAGAGTGTATTTTGTGAACAACAT 466  
QY 301 GTTGTCAATTTGCTTTCATTCATTTATTTGTTACCAAGTTGATGCTGATATTCAGTACAA 360  
|||||  
Db 467 GTTGTCAATTTGCTTTCATTCATTTATTTGTTACCAAGTTGATGCTGATATTCAGTACAA 526  
QY 361 TGCTTTTACATGGAAGCTGATTAATAACAGTTAGTGCACAGTTAGAGTATCTGAAATACA 420  
|||||  
Db 527 TGCTTTTACATGGAAGCTGATTAATAACAGTTAGTGCACAGTTAGAGTATCTGAAATACA 586  
QY 421 ACTGCTTTTCAAACTCAAAATTTGCCGATGCCAGTATGCCGTTATGAAATTTGGATGT 480  
|||||  
Db 587 ACTGCTTTTCAAACTCAAAATTTGCCGATGCCAGTATGCCGTTATGAAATTTGGATGT 646  
QY 481 GGACCAACCGGTCACACAGTTCATTTGCTATCATTTGGTCAGCCAGTTTATCATTAATGG 540  
|||||  
Db 647 GGACCAACCGGTCACACAGTTCATTTGCTATCATTTGGTCAGCCAGTTTATCATTAATGG 706  
QY 541 ACATGCGATTCGTAACCGTTGATCTTCTGCGCGTTGCTGCTGCTGCTGCTGCTGCTGCT 600  
|||||  
Db 707 ACATGCGATTCGTAACCGTTGATCTTCTGCGCGTTGCTGCTGCTGCTGCTGCTGCTGCT 766  
QY 601 GATGTAACCGGATATCTGTAAGATTTCTAAATGCTGATGATGATGATGATGATGATGAT 660  
|||||  
Db 767 GATGTAACCGGATATCTGTAAGATTTCTAAATGCTGATGATGATGATGATGATGATGAT 826  
QY 661 TTGCTAAATATTTGGAATATCCAAAGATTTAATGGTCGCCAAGAAGCTCAGTATAC 720  
|||||  
Db 827 TTGCTAAATATTTGGAATATCCAAAGATTTAATGGTCGCCAAGAAGCTCAGTATAC 886  
QY 721 AATATGCGGATGATGATCAGCTTTTCTATCATTTGCCAGATCAGTATTTACCATTAAGAA 780  
|||||  
Db 887 AATATGCGGATGATGATCAGCTTTTCTATCATTTGCCAGATCAGTATTTACCATTAAGAA 946  
QY 781 CCAATATGCGGATGATGATCAGCTTTTCTATCATTTGCCAGATCAGTATTTACCATTAAGAA 840  
|||||  
Db 947 CCAATATGCGGATGATGATCAGCTTTTCTATCATTTGCCAGATCAGTATTTACCATTAAGAA 1006  
QY 841 ACAAGTGTGCGCGACGCAAAACCTGCTGACAGCTGCGCACTGCTTACTCAAGAAAGA 900  
|||||  
Db 1007 ACAAGTGTGCGCGACGCAAAACCTGCTGACAGCTGCGCACTGCTTACTCAAGAAAGA 1066  
QY 901 TCTGCAACACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 960  
|||||  
Db 1067 TCTGCAACACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1126  
QY 961 GATGATATTCAGCTTTGCGCAGTTGATTAACGTCACCGTGCACCTTCTGCAACATTAAGGA 1020  
|||||  
Db 1127 GATGATATTCAGCTTTGCGCAGTTGATTAACGTCACCGTGCACCTTCTGCAACATTAAGGA 1186  
QY 1021 CAACCTGTAATCTTGTGTCAGTACAAATGATGATGATGATGATGATGATGATGATGAT 1080  
|||||  
Db 1187 CAACCTGTAATCTTGTGTCAGTACAAATGATGATGATGATGATGATGATGATGATGAT 1246  
QY 1081 ATGTTTATGCGTTTAAAGCATTCGATTTGCTGCGCTCATTTTAAACATTTGCTTAAA 1140

```

|||||
Db 1247 ATGTTTATGGTTTAAAGCATGATTGATGTCGCGCATTTATACCATTTGTTTAA 1306
OY 1141 TTTCGTCAAAATCAGAAGCA 1161
|||||
Db 1307 TTTCGTCAAAATCAGAAGCA 1327

RESULT 4
AR204519/c 1779 bp DNA linear PAT 20-JUN-2002
LOCUS AR204519 Sequence 2 from patent US 6368600.
DEFINITION AR204519
ACCESSION AR204519
VERSION AR204519.1 GI:21501866
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1779)
AUTHORS Chandrasekar, R. and Morales, T. H.
TITLE Parasitic helminth cuticlin nucleic acid molecules and uses thereof
JOURNAL Patent: US 6368600-A 2 09-APR-2002;
FEATURES
source 1..1779
location/Qualifiers
BASE COUNT 575 a 502 c 327 g 575 t
ORIGIN

Query Match 100.0%; Score 1161; DB 6; Length 1779;
Best Local Similarity 100.0%; Pred. No. 6.3e-262;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANGATGATTCGTTCTTATTCCTTTCGTACTACATTTATGATGCTTATTCGATTCG 60
Db 1613 ATGATGATTCGTTCTTATTCCTTTCGTACTACATTTATGATGCTTATTCGATTCG 1554
OY 61 GTTGCAATGGTGCAGAGTGCAGCAGAAATGATGAGCAACCTTCATTAAC 120
Db 1553 GTTGCAATGGTGCAGAGTGCAGCAGAAATGATGAGCAACCTTCATTAAC 1494
OY 121 AATTTTAATACAGTAAATGATGAGCAAGCATGTTTATGATGAAGTCTTATGATCAA 180
Db 1493 AATTTTAATACAGTAAATGATGAGCAAGCATGTTTATGATGAAGTCTTATGATCAA 1434
OY 181 GAAGTTCCGTAATGATGAGTGAAGTGCATGAGTTGCCGAATTTCACTTCATTTGAT 240
Db 1433 GAAGTTCCGTAATGATGAGTGAAGTGCATGAGTTGCCGAATTTCACTTCATTTGAT 1374
OY 241 TCATGCAATGTCGGGTGCAGATGCTGAAATCCACGATGATTTTGTAAACAACACT 300
Db 1373 TCATGCAATGTCGGGTGCAGATGCTGAAATCCACGATGATTTTGTAAACAACACT 1314
OY 301 GTTGCAATTCGTTTCATTCATTTATTTGTAAACAAGTTGATGAGTGCATGAGTACA 360
Db 1313 GTTGCAATTCGTTTCATTCATTTATTTGTAAACAAGTTGATGAGTGCATGAGTACA 1254
OY 361 TCGTTTACATGAGAGTGAATAAACAGTTAGTGACAGATTTGAGGTATCTGAATACACA 420
Db 1253 TCGTTTACATGAGAGTGAATAAACAGTTAGTGACAGATTTGAGGTATCTGAATACACA 1194
OY 421 ACTGCTTTTCAACATCAATTTGCCGATGCAGATGAGTGCAGATTTGAGGTATCTGAAT 480
Db 1193 ACTGCTTTTCAACATCAATTTGCCGATGCAGATGAGTGCAGATTTGAGGTATCTGAAT 1134
OY 481 GGACCAACCGGTCAACAGTTCAATTTGATGATGTCAGCAGATTTATCATATAAATG 540
Db 1133 GGACCAACCGGTCAACAGTTCAATTTGATGATGTCAGCAGATTTATCATATAAATG 1074
OY 541 ACATGCAATTCGAACCGGTGATGATGTCGCGGTTGTCATTCCTGCTTGTGAT 600
Db 1073 ACATGCAATTCGAACCGGTGATGATGTCGCGGTTGTCATTCCTGCTTGTGAT 1014
OY 601 GATGTAACGGTGAATCTGTGAATTTCTAAATGCTGATGATGTCGCTTGTGAATAAT 660

```

```

|||||
Db 1013 GATGTAACGGTGAATCTGTGAATTTCTAAATGCTGATGAGTGTCTGTTGAATAAT 954
OY 661 TTGCTAAATTAATTTGAATATCCACAGATTTTAATGGTGGCCCAAGATCAGTATAC 720
Db 953 TTGCTAAATTAATTTGAATATCCACAGATTTTAATGGTGGCCCAAGATCAGTATAC 894
OY 721 AATATGCGGATCGATCAGAGCTTTTCTATCATAGCAGATCAGTATTAACATTAAGAA 780
Db 893 AATATGCGGATCGATCAGAGCTTTTCTATCATAGCAGATCAGTATTAACATTAAGAA 834
OY 781 CCAATAGCAATGTTTGTGACACCAATGTTGAACACCAAGATTTGCGAGTGTAA 840
Db 833 CCAATAGCAATGTTTGTGACACCAATGTTGAACACCAAGATTTGCGAGTGTAA 774
OY 841 ACAGTGTGCGCCAGACCAAAACCTGCTGAGCTGCGCACTTCGTTACTCAAGAAAGA 900
Db 773 ACAGTGTGCGCCAGACCAAAACCTGCTGAGCTGCGCACTTCGTTACTCAAGAAAGA 714
OY 901 TCTGCAGAACCGGAGAAATATCATTTGATGATGATGATGATGATGATGATGATGAT 960
Db 713 TCTGCAGAACCGGAGAAATATCATTTGATGATGATGATGATGATGATGATGATGATGAT 654
OY 961 GATGATTAATCAAGCTTTGGCAGTGTGATTTACGTCACCGTGCACTTGTGCAATATGA 1020
Db 653 GATGATTAATCAAGCTTTGGCAGTGTGATTTACGTCACCGTGCACTTGTGCAATATGA 594
OY 1021 CAACCTGTAATACCTGCTCAGTACCAAAATGATGATGATGATGATGATGATGATGAT 1080
Db 593 CAACCTGTAATACCTGCTCAGTACCAAAATGATGATGATGATGATGATGATGATGATGAT 534
OY 1081 ATGTTTATGGTTTAAAGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 533 ATGTTTATGGTTTAAAGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 474
OY 1141 TTTCGTCAAAATCAGAAGCA 1161
Db 473 TTTCGTCAAAATCAGAAGCA 453

```

```

RESULT 5
AF125580 1380 bp mRNA linear INV 01-AUG-2000
LOCUS AF125580 Wuchereria bancrofti cuticulin-1 (cut-1) mRNA, complete cds.
DEFINITION AF125580
ACCESSION AF125580
VERSION AF125580.2 GI:9625063
KEYWORDS
SOURCE
ORGANISM
Wuchereria bancrofti.
Wuchereria bancrofti.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Wuchereria.
REFERENCE
1 (bases 1 to 1380)
Ramzy, R., Helmy, H., Adely, M., Curtis, K. and Well, G.
Wuchereria bancrofti L3 cuticulin-1 cDNA
unpublished
2 (bases 1 to 1380)
Ramzy, R., Helmy, H., Adely, M., Curtis, K. and Well, G.
Direct Submission
Submitted (03-FEB-1999) Research & Training Center on Vectors of
Diseases, Ain Shams University, Abassia Square, Cairo, Egypt
3 (bases 1 to 1380)
Curtis, K., Ramzy, R., Helmy, H., Adely, M. and Well, G.
Direct Submission
Submitted (15-JUL-1999) Infectious Diseases, Washington University
School of Medicine, 216 S. Kings Highway, St. Louis, MO 63110, USA
Sequence update by submitter
On Aug 1, 2000 this sequence version replaced gi:4741874.
REMARK
COMMENT
FEATURES
location/Qualifiers
source
1..1380
/organism="Wuchereria bancrofti"
/db_xref="taxon:6293"
/dev_stage="L3 larvae"
/country="Egypt; Nile Delta"

```

gene	1.	1.380
/gene="cut-1"	108.	1.280
/gene="cut-1"		
/note="Similar to Ascaris lumbricoides cuticulin-1"		
/codon_start=1		
/product="cuticulin-1"		
/protein_id="AAD28743.2"		
/db_xref="GI:9625064"		
/translation="MLRLISLIVAFITIVAIPIVDNGVEGEPEIEGPTSTYENFR NPFEHAYVVGSLVDGCGRNDGEGGAVLEIPDSSNVARHRSINRGIFVITTVNR SFGHLYPTKIDRAVRYNOCFYMEADKITYSTOLEVSEITATQIVPAPVRCRYELLVG PSGQVPVPAITGDPVYHKWCDSEIVTFCAYVHSCVVDGNGDKVALLNADSCALDK FLUNLEIPDILMAGOEAHYKTYADRSQLEYYQOISTILEPNSECPOCPTGOGG AIRKRVGAALLRQAAAFRYLAKRKRDVDENIVDRTDLIDINEETSLSPNLRHRS SLAHNGEHPIYATVMTQOIGSLTGTGLMGLFVLSVATVIAITLLRSHSTKV"	456 a	224 c 244 g 454 t 2 others

Query Match	55.78;	Score	646.4;	DB	3	Length	1380;
Best Local Similarity	74.48;	Pred	No 3.8e-141;				
Matches	844;	Conservative	0;	Mismatches	281;	Indels	9;
						Gaps	2;

OY	1	ANGAGTATCGCTCATTTGCTTCGTACTACACTATTATGGATTGCTTATTCGATCCG	60
Db	105	ATTATGCTTCGATTAAATAGCATATTAGTAGCATTATTACATATGTAATCCATATCCG	164
OY	61	GTTGCACAATGGTGGAAGGTGACGCCAGAAAATGGAATGGAGCAACACTCAATTAACATC	120
Db	165	GTTGATAATGGTGTGGAAAGGTGACACAGAATAATGGAATGGTGTCCAACATCAATTACGGTT	224
OY	121	AATTTAATACACGTAATGCAATTCGAAGCACATGTTTATGTGAAGGTCTTATGATCAA	180
Db	225	AATTTTAACACTCGAAATCCATTGTAAGACATGTTTATGTAAAGATATATGATCAA	284
OY	181	GAAGGTTCCTCCGTAATGATGAAGGTGAGGTGACGTCAAGTTGGCCGAATTCACCTCCATTGAT	240
Db	285	CAAGGATGTCCGTAATGATGAAGGTGAGGTGACGACAGGATGCTGGAAATGAGCATACCATTGGAT	344
OY	241	TGATGCAATGTGTGCGCGTAPACAGATCTCTGAATCCAGCTGTGATATTTTGTAAACAACACT	300
Db	345	TCATGTAATGTATGAGCAGCTACCGTCTGTTTAATCCAGCTGGAATATTGTGCACACACACA	404
OY	301	GTTGTCATTGTGGTTTCATCCATTAATTTGTTCOCAAAGTTGATCCGTGCATTCGAGTACAA	360
Db	405	GTTGTCATTTTATTTTCATTCACATTAATTTGTACAAAAAGTTGATCGGTCATTCGTGTACAA	464
OY	361	TGCTTTTACATGGAAGCTGATAAACAAGTGTAGTCGACAGATTTGAGATGTGCAATTCACA	420
Db	465	TGCTTTTACATGGAAGCTGATAAACGTGTGTAGCACACAAATTTGAAGTTTCGGAATTTACA	524
OY	421	ACTGCTTTCCAAATCCAAATTTGTGCCGATGCCAGTATGCCGTTATGAATTTGGATGGT	480
Db	525	ACACGATTTGGACAACCAAAATTTGTACCAATGCGCTGTATGTCATTAAGATTTCTTGATGGT	584
OY	481	GGACCAACCGGTGCAACCGAGTTTCTATNCATGTGTGACGCGAGTTTATCATPAAAGG	540
Db	585	GGACGATCAGAACCAACCGGTTCATTTGCTACTCCATTTGAGACAACCGGTTTATCATPAAAGG	644
OY	541	ACATGCAATTCGAAACCGTTGATACTTCTGCGCGGATGTCCATTCCTGCTTGTGTCGAT	600
Db	645	ACATGCTGATTTGGAAACAGTGTATACATTTCTGTGGGGGTGATACATTTCAATGCTTTGGTAT	704
OY	601	GATGCTAACCGGTGATCTGCTGTGGAATTTCTAAATGCTGATGATGTGCTCTTGATPAAATAT	660
Db	705	GATGCTAATGGTGTAAAGTAGACCTATTAAATCTGTAGTGGTGTGTCTTTGGACAAAAATTT	764
OY	661	TTCCTAATAATTTTGGATATCCAAACAGATTTAATGGCTGCGCAGAAGATCCACGATATAC	720
Db	765	CTACTATATAATTTTAGAATATCCAAACGATTTTAATGGCTGACCAAGACACACGATATAC	824
OY	721	AAATATGGGATGCATACACGCTTTTCTATCATATGCCAGATCAGTATTACCATTAAAGAA	780

Db	825	AAATAGCGTGATGCATCACAACCTTTTTTTATGCATATGTCACAAATTAGCATTAACAAATTAAAGAA	884
QY	781	CCAAATAGCGGAATGTGTTTGCACCAANGTTCAGAACCAAGAAGATTCGGAGCTGTTAAA	840
Db	885	CCAAATAGGAATGTTCACAGCACCACAAATGCATGAAACCCACAGGTTTTGGCCAATTAAA	944
QY	841	ACAGG---TGGTGCCGACAGCAAAACCTGCTGCAGCTGCGCAACTTCGTGTTACTCAAGAAA	897
Db	945	ACGCCCAATGGTGTCACACTGCATTTACGACAGCGTGCACATTTAGATATTTGAAAAAG	1004
QY	898	AGATGTCCAGAACCCGAGATATCATTTGATGTACGAAACGATATCAACACCCCTGGAATT	957
Db	1005	AGAGATGTGAAGATGAAGAAATATTTGTTGATGTACGACCGAGATCTTAATGTCAGTATTA	1064
QY	958	AGCGATGATTAATCAAGCTTTGCCAGTTGATTAACGTCAACCGTCGCACTCTG-----CAA	1011
Db	1065	AATGAAGATTTACCTCATTACCAATAGGCGGTACGTCATCGATCATTCATTTGTTAGCAGCT	1124
QY	1012	CATATGGACCAACCTGTATTACTTCTGTGACGTACAAATATGATATCGATGTCACCAATT	1071
Db	1125	GAAATATGGTCATCCCGTATATGTTTCAACAAAGACACAAAGATATATGATGTCAATTAAC	1184
QY	1072	GGCCTCCTCAATGTTTATGGGTTTAGCAATTCGATGATGCTGCGCGTCAATT	1125
Db	1185	GGTTTACACTGCCGATATGCTAATATTTGTTATGTTGTTGTTGCTACAAATT	1238

[illegible]

Query Match	43.4%	Score 504.2	DB 6	Length 1372
Best Local Similarity	76.9%	Pred. No. 9e+108		
Matches 652; Conservative	0	Mismatches 193	Indels 3	Gaps 3

QY	1	ATGATGATTCGCTCTTATTGCTTTCTGTACTACACTTAATG-CATTGCTCTATTGCATTCG	59
Db	32	ATTATGCAATTAATCTTCCTGTTATTCTCTACTTTCTACTTACCTACAGTATCATGATGCC	91
QY	60	GGTGACAATAGTGTCGCAAGGTGAGCCGGAATTAAGTAAGTGGACCAACTTCATAACAT	119
Db	92	TATTGACAAATGCTTCGAAGTGAACCTGAAATAGAAATGTCGGCAGGTTGATPAACAT	151
QY	120	CAATTTTAATACACGTAAATGCAATTCGAAAGCATGTTATGTGAAGGTCTTATGATCA	179
Db	152	CAATTCAAATCACTGAATAATACATTTGGAAGGACAGCTATGTAAAGSACCTATGATCA	211
QY	180	AGAAGTGTCCGTAATATGTAAGSTGGACGTCGAAGTGGCGGAATTTCACTTCATTTGA	239
Db	212	GGAATGAATGTCGTCGAATAGTAATGAGCGGACAGGTAGCTGGAATCGAATGGCAATGGA	271
QY	240	TTCATGCAATGTTGCGGCTACACGATCTCGAATCCAGCGTGTATTTTGTGAACAACAC	299
Db	272	TGTGTGTAATGTGAACGATCAGCATCCTTAAACCTCTCGTGGTCTTTTGTGAACAACGT	331
QY	300	TGTTGTCATTTCCGTTTATCCCATTTATTTGTGAACCAAGTGAATGTCATATCGAGTGA	359

Db 332 AGTTGCAATTTACATTCACAAAATTCGTTACAAAATAGATCGACATATTCGTATACA 391  
QY 360 ATGCTTTTACATGG-AAGCTGATAAAAACAGTTAGTGC-ACAGATTGAGGATATCGAAATC 417  
Db 392 ATGTTTTTATATGGAAGAGTGAATAGACCGTTAGTACTGCTCTTGAAGATATCGAAATC 451  
QY 418 ACAATGCTTTTCAAACTCAAAATTTCCGATGCCAGTATCCGTTATGAAATTTTGAT 477  
Db 452 ACTACAGCATTCACAACTCAAGTGCATCAATGCCGATATGCGATATGAGATTTTGAA 511  
QY 478 GGTGGACCAACCGGATCAACAGTTCAATTTGCTATCATTTGGTGGAGCGATTTTATCAA 537  
Db 512 GGTGGACCAACGTTGTCACCTGTTGCAATTTGCAATATGGAGATCATATATATCAAA 571  
QY 538 TGGACATGCGATTTCTGAAACCGTTGATCTTCTGCGGCTGTCTCCATTCCTGCTTTGTC 597  
Db 572 TGGACATGATATCAAGACATACAGATATCTGTGCAATTTAGTATCATTCATGTTTG 631  
QY 598 GATGATGTTAGCGGTATCTGTGGAATTTCTAATGCTGATGATGCTCTTGTATTA 657  
Db 632 GATGATGGAAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 691  
QY 658 TATTTCTAATATATTTGGAATATCCACAGATTTAATGCTGGCCAGAGCTCACTGA 717  
Db 692 TATTTCTAATATATTTGGAATATATACAGATTTAATGCTGGCCAGAGCTCACTGT 751  
QY 718 TACAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
Db 752 TATTAATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 811  
QY 778 GAACCAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837  
Db 812 GAGCCACATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 871  
QY 838 AAAACAG 845  
Db 872 AATCTGG 879

RESULT 7  
AR204523/c 1372 bp DNA linear PAT 20-JUN-2002  
LOCUS AR204523  
DEFINITION Sequence 7 from patent US 6368600.  
ACCESSION AR204523  
VERSION AR204523.1 GI:21501871  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1372)  
AUTHORS Chandrasekar, R. and Morales, T.H.  
TITLE Parasitic helminth cuticulin nucleic acid molecules and uses thereof  
JOURNAL Patent: US 6368600-A 7 09-APR-2002;  
FEATURES  
source 1.1372  
BASE COUNT 406 a 281 c 229 g 456 t  
ORIGIN

Query Match 43.4%; Score 504.2; DB 6; Length 1372;  
Best Local Similarity 76.9%; Pred. No. 9e-108;  
Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;

QY 1 ATGATGATGCTGCTTATGCTTTCTGATCTATGATCTATG-CATTTGCTTATTCGATTC 59  
Db 1341 ATATCATATATCTCTCTTATCTTATCTTCTACTTTCTACTGACTACGATTCATGCC 1282  
QY 60 GGTGCAATGCTGCTGCAAGTGCAGCAATGATGATGATGATGATGATGATGATGAT 119  
Db 1281 TATGCAATGCTGCTGCAAGTGCAGCAATGATGATGATGATGATGATGATGATGAT 1222  
QY 120 CAATTTATACAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 179

Db 1221 CAATTTCAATAGTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1162  
QY 180 AGAAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239  
Db 1161 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1102  
QY 240 TTGATCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299  
Db 1101 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042  
QY 300 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359  
Db 1041 AGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 982  
QY 360 ATGCTTTTACATGG-AAGCTGATAAAAACAGTTAGTGC-ACAGATTGAGGATATCGAAATC 417  
Db 981 ATGTTTTTATATGGAAGAGTGAATAGACCGTTAGTACTGCTCTTGAAGATATCGAAATC 922  
QY 418 ACAATGCTTTTCAAACTCAAAATTTCCGATGCCAGTATCCGTTATGAAATTTTGAT 477  
Db 921 ACTACAGCATTCACAACTCAAGTGCATCAATGCCGATATGCGATATGAGATTTTGAA 862  
QY 478 GGTGGACCAACCGGATCAACAGTTCAATTTGCTATCATTTGGTGGAGCGATTTTATCAA 537  
Db 861 GGTGGACCAACGTTGTCACCTGTTGCAATTTGCAATATGGAGATCATATATCAAA 802  
QY 538 TGGACATGCGATTTCTGAAACCGTTGATCTTCTGCGGCTGTCTCCATTCCTGCTTTGTC 597  
Db 801 TGGACATGATATCAAGACATACAGATATCTGTGCAATTTAGTATCATTCATGTTTG 742  
QY 598 GATGATGTTAGCGGTATCTGTGGAATTTCTAATGCTGATGATGCTCTTGTATTA 657  
Db 741 GATGATGGAAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 682  
QY 658 TATTTCTAATATATTTGGAATATCCACAGATTTAATGCTGGCCAGAGCTCACTGA 717  
Db 681 TATTTCTAATATATTTGGAATATATACAGATTTAATGCTGGCCAGAGCTCACTGT 742  
QY 718 TACAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
Db 621 TATTAATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 562  
QY 778 GAACCAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837  
Db 561 GAGCCACATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502  
QY 838 AAAACAG 845  
Db 501 AATCTGG 494

RESULT 8  
AR204531 892 bp DNA linear PAT 20-JUN-2002  
LOCUS AR204531  
DEFINITION Sequence 16 from patent US 6368600.  
ACCESSION AR204531  
VERSION AR204531.1 GI:21501881  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 892)  
AUTHORS Chandrasekar, R. and Morales, T.H.  
TITLE Parasitic helminth cuticulin nucleic acid molecules and uses thereof  
JOURNAL Patent: US 6368600-A 16 09-APR-2002;  
FEATURES  
source 1.892  
BASE COUNT 290 a 142 c 168 g 292 t  
ORIGIN

Query Match 39.7%; Score 461; DB 6; Length 892;  
Best Local Similarity 78.1%; Pred. No. 1.3e-97;



Query Match	39.7%;	Score 461;	DB 6;	Length 892;
Best Local Similarity	78.1%;	Pred. No. 1.3e-97;		

RESULT 10	CEP22B5/C	25823 bp	DNA	linear	INV 23-JAN-2002
LOCUS	CEP22B5				
DEFINITION	Caenorhabditis elegans cosmid F22B5, complete sequence.				
ACCESSION	Z50044				
VERSION	Z50044.1				
KEYWORDS	HTG; Cuticulin; Elongation factor; GTP-binding ADP-ribosylation factor; Phenylalanyl-tRNA synthetase; RNA binding.				
SOURCE	Caenorhabditis elegans.				
ORGANISM	Caenorhabditis elegans.				
REFERENCE	1				
AUTHORS	none.				
TITLE	Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium				
JOURNAL	Science 282 (5396), 2012-2018 (1998)				
MEDLINE	99069613				
PUBMED	9851916				
REMARK	The <i>C. elegans</i> Sequencing Consortium.				
REFERENCE	2 (bases 1 to 25823)				
AUTHORS	Sims, M.A.				

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (13-JUL-1995) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@anger.ac.uk or r@nemastode.wustl.edu	Coding sequences below are predicted from computer analysis, using predictions from GeneFINDER (P. Green, U. Washington), and other available information.
gene	Current sequence finishing criteria for the <i>C. elegans</i> genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.	IMPORTANT: This sequence is not the entire insert of clone F22B5. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
CDS	The true left end of clone F22B5 is at 24607 in sequence Z66522.	The true right end of clone F22B5 is at 16002 in sequence Z66523.
gene	The true left end of clone M0506 is at 25723 in this sequence. The true right end of clone F14R5 is at 4610 in this sequence. The start of this sequence (1..101) overlaps with the end of sequence Z66522.	The end of this sequence (25723..25823) overlaps with the start of sequence Z66523.
CDS	For a graphical representation of this sequence and its analysis see: <a href="http://wormbase.sanger.ac.uk/pel/ace/elegans/seq/sequence?name=F22B5">http://wormbase.sanger.ac.uk/pel/ace/elegans/seq/sequence?name=F22B5</a>	IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
FEATURES	Location/Qualifiers	
source	1..25823	
gene	/organism="Caenorhabditis elegans"	
CDS	/db_xref="taxon:6239"	
gene	/chromosome="II"	
CDS	/clone="F22B5"	
gene	join(411..462,584..663,745..1127,1216..1471)	
CDS	/gene="F22B5.2"	
gene	join(411..462,584..663,745..1127,1216..1471)	
CDS	/gene="F22B5.2"	
gene	/note="contains similarity to Pfam domain: PF00076 (RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)), Score=95.9, E-value=2.6e-25, N=1	
CDS	CDNA EST yk39048.5 comes from this gene	
gene	CDNA EST yk49946.3 comes from this gene	
CDS	CDNA EST yk39048.3 comes from this gene	
gene	CDNA EST yk49946.5 comes from this gene"	
CDS	/codon_start=1	
gene	/protein_id="CAA90354.1"	
CDS	/db_xref="GI:3876226"	
gene	/db_xref="SWISS-PROT:Q19706"	
CDS	/translation="MAPAPVAVSWAEAEQDNAPHIQSGADGTRTETAFTEGVNRK VVQKRVINKRPKVVADRRKRVKSGSGEPAPVATTVAAEEVDQAFRRNGEG ILVDQEDKOTAKTTSREHCRHCKGNDHNSCTGPKVMTQLEADADKDTGRRAGG MRPDGROIDRNSDENTCRVNLPOENNEDELDLDFGKIGRIVIRFIARDVVTGLPKG FAFVFEESRDAAARAIALINDIRMYHMYLKYEWRPSN"	
gene	join(1870..2158,2209..2339,2534..2668)	
CDS	/gene="F22B5.1"	
gene	join(1870..2158,2209..2339,2534..2668)	
CDS	/gene="F22B5.1"	
gene	/note="contains similarity to Pfam domain: PF00025 (ADP-ribosylation factor family), Score=295.0, E-value=2.9e-85, N=1"	
CDS	/codon_start=1	
gene	/protein_id="CAA90353.2"	
CDS	/db_xref="GI:14530408"	
gene	/db_xref="SWISS-PROT:Q19705"	
CDS	/translation="MGFLKILRKORARERERILLIGDNAGKTTMKKFLIDETPTDI EPTIGFDIKYVHFADQNLMDVGGQSKRSYWKMYTFESTDALIIVVDSSDRELRLOC	
gene	SEELKKLGEERLAGASLLVLANKSDLPGAIDVNSIAQVLDLHSIKSHHWLIFSCAL	
CDS	SGDRIVQATWMTKCDVGSRLFTLD"	
gene	complement(join(2833..3042,3088..3700,3823..4048,4273..4393))	
CDS	/gene="F22B5.3"	
gene	complement(join(2833..3042,3088..3700,3823..4048,4273..4393))	
CDS	/gene="F22B5.3"	
gene	/note="similar to cuticulin	
CDS	CDNA EST EMBL:701970 comes from this gene	
gene	CDNA EST CEBSX90RB comes from this gene	
CDS	CDNA EST yk38667.3 comes from this gene	
gene	CDNA EST yk38667.5 comes from this gene	
CDS	CDNA EST yk504b12.3 comes from this gene	
gene	CDNA EST yk54392.3 comes from this gene	
CDS	CDNA EST yk52b8.3 comes from this gene	
gene	CDNA EST yk504b12.5 comes from this gene	
CDS	CDNA EST yk54392.5 comes from this gene	
gene	CDNA EST yk551h11.5 comes from this gene	
CDS	CDNA EST yk590c10.5 comes from this gene	
gene	CDNA EST yk652b8.5 comes from this gene"	
CDS	/codon_start=1	
gene	/protein_id="CAA90355.1"	
CDS	/db_xref="GI:3876227"	
gene	/db_xref="SPTREMBL:Q19707"	
CDS	/translation="MARSLSGLGLDLVASVAPVNNVNGEEVEGCPSTVYNN TRNATGEVTVKGLFDQOECRDEBGRVAGIELEFDDCNVARTSLNPKGVFTTV VSEFHPQVTVYDRAVRYVQCEYMEADKVSQIEVSDLTTFAPQVPMPICKEILN VSGTEPVOFATIIQQVYVHKWTCDESEYDFCAVHSCVVDGNDVQIOLDEAGCAL DKFLNNLEYPTLMAGOGEAHVYKAYADSDYFQCOQIITVKEENEECARPTGCEPOG FGAVYQAOVOTOPFVRLKRSAPVYENILDVRAETLTVELEGNLPSSTQAOALVRS REIGDSRQELCTSSFTISVTVYGLTGVVAFITMYIYSRMAVPSDRQNSAC"	
gene	join(9651..9890,9940..10234)	
CDS	/gene="F22B5.4"	
gene	join(9691..9890,9940..10234)	
CDS	/gene="F22B5.4"	
gene	/note="CDNA EST yk304c8.3 comes from this gene	
CDS	CDNA EST yk304c8.5 comes from this gene	
gene	CDNA EST yk594h3.3 comes from this gene	
CDS	CDNA EST yk594h3.5 comes from this gene"	
gene	/codon_start=1	
CDS	/protein_id="CAA90356.1"	
gene	/db_xref="GI:3876228"	
CDS	/db_xref="SPTREMBL:Q19708"	
gene	/translation="MSLSKARKTKMFRVLSRNASTNSVSPARIQLKPAEAGHPT SRMSRDRFVKVAIQLGDTPTQFLVRLGHAIVYPLFVLTAAEVLFCASASTWSFG KAEIWLDRSNSKAPDWERLDRDTYWKMPVAFDDGRTKRCELMEOLODEMLEAARK RGTB"	
CDS	complement(join(10849..10984,11033..11343,11397..11521,11578..11832,11917..12154))	
gene	complement(join(10849..10984,11033..11343,11397..11521,11578..11832,11917..12154))	
CDS	/gene="F22B5.5"	
gene	complement(join(10849..10984,11033..11343,11397..11521,11578..11832,11917..12154))	
CDS	/gene="F22B5.5"	
gene	/codon_start=1	
CDS	/protein_id="CAA90357.1"	
gene	/db_xref="GI:3876229"	
CDS	/db_xref="SPTREMBL:Q19709"	
gene	/translation="MNCCKQASKAVFVLGVGFNKASSRCFSSKELIPKDLKIVKNPTS EQLKQKRLVGDVNDIKLKEETAKMDNQLVLLCHKAKWSJALACYNNLNSNMSIY SDSVLSGTOSITVKSLSNRSIPDFQFGLSYQPDNTYNLPHLHMEHATDFKSVKMS GGCYVSLAAVWRKVLTKVKGPTVYVSHKPPDEVNDELPPDDVYKFSQVPEHVYK YDNSIFPEYKERYLAKFKNGVGVAAKNGEYVIGLVTBESSGNEICPTICDKN AAQAIRPSTIIDDPMKDLKYLVRCSKFEKDSATWIRPPLRSHEMTPFAHIFNRYI PDGLNFSKVFVSNMSPNAPC"	
CDS	complement(join(12240..12394,12467..12541,12691..12784,13262..13392,13465..13524,13876..14032))	
gene	complement(join(12240..12394,12467..12541,12691..12784,13262..13392,13465..13524,13876..14032))	
CDS	/gene="F22B5.6"	
gene	complement(join(12240..12394,12467..12541,12691..12784,13262..13392,13465..13524,13876..14032))	
CDS	/gene="F22B5.6"	
gene	/codon_start=1	



```

/protein_id="CA90358.1"
/db_xref="GI:3876230"
/ld_xref="SPRMBL:019710"
/translation="MKNPYGTAAYVNSCLSTWISOKRNPSEYLBALIPVPESEMWL
LAQYHIHESSENITISWNMLNNSRKGIVNANNTNHEGLCVTAFAICITSLIIVDIL
FKHFTFRYVNVNIAVAGCLFVFTVGLLHLEIWFAMSNITSCROGLDITLITS
NTESEVERKTESAGCEYTAISSEKKNVFPDLQVAFIPHLILTFPSWTFSLYPRW
HEDFE"
gene
join(15384..15575,15622..16479,16527..17105,18608..18694
18794..18946,18996..19449,19497..19843,20042..20554,
20672..21600,21677..21812)
/feature="F22B5.7"
join(15384..15575,15622..16479,16527..17105,18608..18694
18794..18946,18996..19449,19497..19843,20042..20554,
20672..21600,21677..21812)
/gene="F22B5.7"
/note="limited similarity with the elongation factor EF-3
(Swiss Prot accession number P25997)
CDNA EST EMBL:M89075 comes from this gene
CDNA EST yk28d8.3 comes from this gene
CDNA EST yk28d8.5 comes from this gene
CDNA EST yk153b7.5 comes from this gene
CDNA EST yk349b7.5 comes from this gene
CDNA EST yk420b2.5 comes from this gene
CDNA EST yk421d3.3 comes from this gene
CDNA EST yk421d3.5 comes from this gene
CDNA EST yk396b3.3 comes from this gene
CDNA EST yk396b3.5 comes from this gene
CDNA EST yk359d5.3 comes from this gene
CDNA EST yk359d5.5 comes from this gene
CDNA EST yk351a5.3 comes from this gene
CDNA EST yk351a5.5 comes from this gene
CDNA EST yk355a11.3 comes from this gene
CDNA EST yk355a11.5 comes from this gene
Query Match 29.2%; Score 339.2; DB 3; Length 25823;
Best Local Similarity 63.3%; Pred. No. 4,5e-69;
Matches 633; Conservative 0; Mismatches 233; Indels 134; Gaps 2;

```

Db	3571	TAACGAGAGACCACATGGSAAACACAGTTCATTTCCGACACTATTGGACAACCAAGTCTATCA	3511
Qy	534	TAAAATGCAATGSCGATTTCTGAAACCGTTGATCTTCTCGCGGGTGTCCATTCTCGT	593
Db	3511	CAAAATGGACCTGTGTTTTTTGTGAACCGTTGATCTTCTCGCGGTGTCTCACCTTGAC	3452
Qy	594	TCTCGATATGGTAAACGGTGTACTGTGGAAATCTTAAATGCGTAGATATGTCCTTGA	653
Db	3451	TCTGTATGTATGGAATATGGTGTATACCGTGCAAAATTTGGATGAGACACGGTTGGCTTGA	3392
Qy	654	TAAATATTTGCTAAATTAATTTTGGAAATATCCACAGATTTTAATGGCTGGCCCAAGACTCA	713
Db	3391	TAAAGTCTTGCTCAACAATTTGGAAATCCACATGATTTGATGGCTGGACAGAGGCTCA	3332
Qy	714	CGTATCCAAATATCGCGATCGATACAGCTTTTCTATCAATCCACAGATCATGATTACAT	773
Db	3331	CGTCAACAATACCGCTGATGCTTCTCACTTTTCTATCAATGCGCAAGATTTCCATTACGT	3272
Qy	774	TAAAGACCAATATAGCAATATGTTTGGACCACAAATGTTGAGACCAAGAATTTGGAGC	833
Db	3271	TAAAGAACCAATATAGGAGATGTGCTCGCCCAACGTCTCTGAGCCACAAGATTTGGAGC	3212
Qy	834	TGTTAAACACAGGTGTGTCGCCACGCAAAACCTGTGAGCTGGGCAACTTCGTTTACTAA	893
Db	3211	AGTTAA-----GCAGCTTACCAAACTGTCTAGTTCTTTCAGAGTTTGGAAACA	3164
Qy	894	GAAATGATTCGACAGACCGGAGAAATATCATTTGATGACAACTGATATCAACACCTTGA	953
Db	3163	GAGATCGGCTCCAGTATGAGAAATATTCCTTGTATGTCGCTGTGAGTCAACAACATTGGA	3104
Qy	954	AATTAGCATGATTAATCAACCTTTGGCACTTGATTTAGCT	993
Db	3103	AGTTTGGAGGAATGTAGTGTCTTCAAAATGSGTGTAT	3064

RESULT 11							
CEC47G2							
LOCUS	CEC47G2	34831 bp	DNA	linear INV 24-JAN-2002			
DEFINITION	Caenorhabditis elegans cosmid C47G2, complete sequence.						
ACCESSION	249125						
VERSION	249125.1	GI:790365					
KEYWORDS	HTG; Cuticulin; Fork head.						
SOURCE	Caenorhabditis elegans.						
ORGANISM	Caenorhabditis elegans						
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;						
	Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.						
REFERENCE	1	none.					
AUTHORS							
TITLE	Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium						
JOURNAL	Science 282 (5396), 2012-2018 (1998)						
MEDLINE	99063613						
PubMed	9851916						
REMARK	The <i>C.elegans</i> Sequencing Consortium.						
REFERENCE	2	(bases 1 to 34831)					
AUTHORS	Palmer,S.						
TITLE	Direct Submission						
JOURNAL	Submitted (27-APR-1995) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jcs@sanger.ac.uk or rwnematode.wustl.edu						
COMMENT	Coding sequences below are predicted from computer analysis, using predictions from GeneFINDER (P. Green, U. Washington), and other available information.						
	Current sequence finishing criteria for the <i>C. elegans</i> genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.						
	IMPORTANT: This sequence is not the entire insert of clone C47G2. It may be shorter because we only sequence overlapping sections once, or longer because we only sequence for a small overlap between neighbouring submissions.						

FEATURES  
source

The true left end of clone C47G2 is at 1 in this sequence. The true right end of clone C47G2 is at 6388 in sequence 249912.  
The true left end of clone T24P1 is at 34731 in this sequence. The true right end of clone T05B9 is at 12750 in this sequence. The start of this sequence (1..101) overlaps with the end of sequence 249129.  
The end of this sequence (34731..34831) overlaps with the start of sequence 249912.  
For a graphical representation of this sequence and its analysis see: - <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=C47G2>  
IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.  
Location/Qualifiers  
1..34831  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="II"  
/clone="C47G2"  
join(2562..2679,2992..3217,3898..4636,4687..4878)  
/gene="C47G2.1"  
join(2562..2679,2992..3217,3898..4636,4687..4878)  
/gene="C47G2.1"  
join(2562..2679,2992..3217,3898..4636,4687..4878)  
/gene="C47G2.1"  
/note="similar to cuticulin  
CDNA EST YK146a2.5 comes from this gene  
CDNA EST YK146a2.3 comes from this gene  
CDNA EST YK262c12.3 comes from this gene  
CDNA EST YK262c12.5 comes from this gene  
CDNA EST YK265b3.5 comes from this gene  
CDNA EST YK27898.3 comes from this gene  
CDNA EST YK27898.5 comes from this gene  
CDNA EST YK439e8.3 comes from this gene  
CDNA EST YK439e8.5 comes from this gene  
CDNA EST YK380a11.3 comes from this gene  
CDNA EST YK583c11.3 comes from this gene  
CDNA EST YK380a11.5 comes from this gene  
CDNA EST YK583c11.5 comes from this gene"  
/codon\_start=1  
/protein\_id="CA88934.1"  
/db\_xref="GI:3875030"  
/db\_xref="SWISS-PROT:Q03755"  
/translation="MTMKPILICLAIVLASATPVNNGEPEVECGNSTIVNT  
RNFPEHYVTKGIDYDAGCRSDGSGQVAGICILPREPCMTAPRSLNPGVSTTV  
ISFHPQVTKVDPAIRIQCFHESDKTVSTQIETVSUTLTAFTQVYPMVCKEITIDG  
KPSGQPIQFATIQOYVHKMTCDSETTDFCAVHSCIVDNGDVQLNEGCLAD  
KFLNNLEPTDLMAGOEAHVYKYADRSQLFYOCQISITIKDGSFCARPTCEPQF  
GAVKQAGAGAHAAAPQAQAVEVQAAPVAAAAPVAAAAPVAAAPVAPRATLAQRL  
LRKRSFEGEGILDYREINTLIDMEGASPSAPEAALVSESVRRRATSTGICILP  
IGFASPLIGTIVATLSATIEFVARPTSHKH"  
join(15433..15770,16205..16531,16772..17108)  
/gene="C47G2.2"  
join(15433..15770,16205..16531,16772..17108)  
/gene="C47G2.2"  
/note="contains similarity to Pfam domain: PF00250 (Fork head domain). Score=22.2, E-value=2.5e-63, N=1  
CDNA EST EMBL:T01618 comes from this gene  
CDNA EST YK123h2.5 comes from this gene  
CDNA EST YK123h2.3 comes from this gene  
CDNA EST YK268f8.3 comes from this gene  
CDNA EST YK268f8.5 comes from this gene  
CDNA EST YK474g11.5 comes from this gene  
CDNA EST YK637f8.3 comes from this gene  
CDNA EST YK670a12.3 comes from this gene  
CDNA EST YK637f8.5 comes from this gene  
CDNA EST YK670a12.5 comes from this gene"  
/codon\_start=1  
/protein\_id="CA88935.1"  
/db\_xref="GI:3875031"  
/db\_xref="SPTREMBL:O18694"  
/translation="MLFSMESILSTKPLEPPPKLEPVTINQVYVDLFRSNTRISE

PTSAVLEHDLKFGESRRRSILGDEPTEDEDGVPYRKANKRNHSTSAADSSDDA  
KODDDDDSTSRKSMGSRKSSHAKPPYATIALAMSLNPEKLTISECEITIK  
FEYKKEKPAWQNSIRHLSLNDCTFVYARGPGPGGNGNVALDPNCEMDNCSFIR  
RERYKKEKNSDYHEHMSHHPPEPPLQPGKPPRNHHPANLPIGLHPNPAVYN  
MPAFIIPQNDISQILSMASRIMPMQAPVSGQKRTSSSSSPNENSSAVSDKLSA  
complement(join(22244..22507,23220..23597))  
/gene="C47G2.3"  
complement(join(22244..22507,23220..23597))  
/gene="C47G2.3"  
complement(join(22244..22507,23220..23597))  
/gene="C47G2.3"  
/codon\_start=1  
/protein\_id="CAB97233.1"  
/db\_xref="GI:9367121"  
/db\_xref="SPTREMBL:Q9A09"  
/translation="MALASSNATSWIEELFGNPFPRKERVVEEKPREYPSAYQMI  
DOMGIKRPNNPPTPIKPIOMLITLPMSEERIMQMGEMCTAKASISGLVGVA  
FAGFLTASVDPQSLMVGDDPTKOLITKOTKREKSSRKSYGRKFGSLGMFSGTECA  
LETIRAKSDWRNGYSGGIVGLGLGRNGISAPYAWGAAPVPESTIIDHYKRC"  
complement(join(24417..24975,25043..25176,25237..25522,  
25668..25907,25967..26054,26106..26227,26277..26608,  
26681..26785,26846..26914))  
/gene="C47G2.4"  
complement(join(24417..24975,25043..25176,25237..25522,  
25668..25907,25967..26054,26106..26227,26277..26608,  
26681..26785,26846..26914))  
/gene="C47G2.4"  
complement(join(24417..24975,25043..25176,25237..25522,  
25668..25907,25967..26054,26106..26227,26277..26608,  
26681..26785,26846..26914))  
/note="similarity to Trichostrongylus colubriformis 11 kd  
secretory protein (Swiss Prot accession number P21937)  
CDNA EST YK29g1.3 comes from this gene  
CDNA EST YK29g1.5 comes from this gene  
CDNA EST YK6d5.5 comes from this gene  
CDNA EST YK179h8.5 comes from this gene  
CDNA EST YK179h8.3 comes from this gene  
CDNA EST YK358b10.3 comes from this gene  
CDNA EST YK358b10.5 comes from this gene  
CDNA EST YK301f4.3 comes from this gene  
CDNA EST YK301f4.5 comes from this gene  
CDNA EST YK295f2.3 comes from this gene  
CDNA EST YK295f2.5 comes from this gene  
CDNA EST YK550d1.3 comes from this gene"  
/codon\_start=1  
/protein\_id="CA88936.1"  
/db\_xref="GI:3875032"  
/db\_xref="SPTREMBL:O18695"  
/translation="MGTISLAVQLTVFELNLSYLKSTIRKONPIYITSTFGHYF  
SLIIVFVLADVAITFPFKCENDRQRIINLTSTPAVPEDELPGGVDPDVLENLM  
RVVYSAGQLLWILIPLOSQYTAGNEFLIGIRAVNNMAYAVISLCELATITVA  
MEKGSINTEENKVIIVSASNTWGLFLLVLLGHLELPSLIMHGNRRLKRTYF  
DIKELASKESEAEENKDIYKRVRLFPNKMNDGSRKRVITILSKRSDVIDNLF  
SHQVIDNMLHDSGCSSEAKELSLIKRTIYANOTLNNATQKVLVDALFLENLAFS  
ESNGYNLEHSRNTQVPIGRFRWYTRLOTPRCRILGIYVPEKTPVLESECTFVSY  
TLSPAFTLEISVSTRFHKYVYQVFAFGLIIVLITSAFTIRLOLTKYIHDPNRHTD  
ENSITFSLILCRLTPICLHNYRGAVALNHDSHSMKSGFIEFQFKLGHLDVETLA  
KQINITYLPICILLCALHYYRGAVALNHDSHSMKSGFIEFQFKLGHLDVETLA  
RISKNSDRQPTQNTNPSFGSSNAGSGTTSFKRSNNKNDERRPMLDEDDDEVEREST  
STKMSPTREPPSSSGFPDDM"  
join(30076..30134,30184..30492,31050..31274,31321..31480,  
31534..31653,31701..31835,31889..32104,32160..32570,  
32639..32935,32998..33215,33263..33651,33802..34049)  
/gene="C47G2.5"  
join(30076..30134,30184..30492,31050..31274,31321..31480,  
31534..31653,31701..31835,31889..32104,32160..32570,  
32639..32935,32998..33215,33263..33651,33802..34049)  
/gene="C47G2.5"  
/note="CDNA EST EMBL:M79633 comes from this gene  
CDNA EST YK179g3.5 comes from this gene  
CDNA EST YK227d8.5 comes from this gene  
CDNA EST YK257e7.5 comes from this gene  
CDNA EST YK319c4.5 comes from this gene  
CDNA EST YK335b12.5 comes from this gene  
CDNA EST YK486d7.5 comes from this gene  
CDNA EST YK513d7.5 comes from this gene  
CDNA EST YK615h7.5 comes from this gene"





Db 200 GCACGGTTGATATAACCGAAGAGAAATCCGAATCTCTCAGCAAT 157

RESULT 15  
AT1173005

GCACGGTTGATATACCGAAGAGAAATCGAATCTTCACCAAT 157

## DEFINITIO

Accession	Size (bp)	Type	Gene
ALU73005	5792	linear	INV 14
Ascaris lumbricoides CUT-1-like cuticlin protein precursor			
(ascut-1) gene, complete cds.			

ACCESSION	073005
VERSION	073005.1
	GI:1657624

**SOURCE**  
**ORGANIS**

Eukaryota; Metazoa;

## REFERENCE

timinoun1, M., and Baz-  
cut-1-1 like genes of

JOURNAL  
MEDICAL

9249070

## REFERENCE AUTHORS

Direct Submission

JOURNAL

Location/Qu

**Sour**

/variety="s  
/db xref="+

gene

join(766. . .

766. .881

misc

/gene="ascu

CDS

/gene="ascu  
/note="cut+1

```

/condon_scar
/product="C

```

/translation

TCOPIQFATIG

HARGÜVLLSE  
882 929

mat

```
join(200 :
/gene="ascu
```

3'UT

**/gene="ascu**

poly

1689 a 1254 c



```

CDS
    /gene="cut-1"
    join(<2..164,266..>484)
    /gene="cut-1"
    /codon_start=1
    /product="cuticlin"
    /protein_id="CA10074.1"
    /db_xref="GI:3858955"
    /db_xref="SPTREMBL:O96775"
    /translation="VISFHPILSTKYDRYRQCFMEADKVTSTQIEVSEITTAQT
    QIVPMFVREIUDGPTQPIQFATIGDPVYHKWTCDESEVDFCAVHSCFVDDGN
    GDVEILSADGALDKYLLNLEYP"
    165..265
    /gene="cut-1"
    /number=1
    /number=1
    /gene="cut-1"
    266..484
    /gene="cut-1"
    /number=2

Intron
    174 a 136 c 134 g 193 t

BASE COUNT
ORIGIN
    174 a 136 c 134 g 193 t

Query Match
    16.4%; Score 190.6; DB 3; Length 637;
    Best Local Similarity 68.3%; Pred. No. 3.7e-34;
    Matches 334; Conservative 0; Mismatches 54; Indels 101; Gaps 1;

    303 TGTCAATTCGTTCAATTCATTTATTTTACCAAGTGTGATCGATATCGAGTACATG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 1 TGTATTTCTTCATTCATTCATTCGTCGACAAAAGTTGATCGACATACGAGTGCAGTG 60

    363 CTTTTCACATGGAAGCTGATGATTAACAGTTAGTGCACAGATTGAGGTATGAAATCACAC 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 61 CTCTATATGGAAGCTGACAAAAGGTCACACCCAGATTGAGGTGCTGAAATCACAC 120

    423 TGGTTTCAACATCAATTCGCCGATGCCAGTATGCCGTTATGAA----- 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 121 TGGTTTCAACATCAATTCGCCGATGCCGTTATGAAAGAGCATGCTTTTGA 180

    469 ----- 468
    Db 181 ATTTTGAATTCATTTATTTATCAATTCAAATCAACATGAAAATCTATTTATCA 240

    469 ----- ATTTGGATGGTGGACCAACCGGTCAACAGTT 501
    Db 241 CATAGTAGATGTGACAAATATTTTATGATTTTGGATGGTGGACCAACCGGACCACTAT 300

    502 CATTTGCTATCAATGCTGTCAGCCAGTTATCATTAATGACATGCCATTCGAAACCGTT 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 301 CAATTTGCTATCAATGCTGTCAGCCAGTTATCATTAATGACATGCCATTCGAAACCGTT 360

    562 GATACCTTTCGCGGTTGTCATTCCTGCTTGTGATGATGATGATGATGATGATGATG 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 361 GATACCTTTCGCGGTTGTCATTCCTGCTTGTGATGATGATGATGATGATGATGATG 420

    622 GAATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 421 GAATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

    682 CCAACAGAT 690
    ||| ||| |||
    Db 481 CCACACAGT 489

RESULT 18
BPA012618 358 bp DNA linear INV 29-JUN-1999
LOCUS BPA012618
DEFINITION Brugia pahangi cut-1 gene, partial.
ACCESSION AJ012618.1 GI:3858956
VERSION AJ012618.1
KEYWORDS cut-1 gene; cuticlin.
SOURCE Brugia pahangi.
ORGANISM Brugia pahangi
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Brugia.
REFERENCE
1 (bases 1 to 358)

```

```

AUTHORS
    Lewis, E., Hunter, S.J., Tetley, L., Nunes, C.P., Bazzicalupo, P. and
    Devaney, E.
TITLE
    cut-1-like genes are present in the filarial nematodes, Brugia
    pahangi and Brugia malayi, and, as in other nematodes, code for
    components of the cuticle
JOURNAL
    Mol. Biochem. Parasitol. 101 (1-2), 173-183 (1999)
MEDLINE
    99339397
PUBMED
    10413052
REFERENCE
    2 (bases 1 to 358)
AUTHORS
    Devaney, E.
TITLE
    Direct Submission
JOURNAL
    Submitted (06-NOV-1998) Devaney E., Veterinary Parasitology,
    University of Glasgow, Bearsden Road, Glasgow, G61 1QH, UK

FEATURES
    source
    1..358
    /organism="Brugia pahangi"
    /db_xref="taxon:6280"
    1..356
    /gene="cut-1"
    join(<1..191,350..>356)
    /gene="cut-1"
    join(<1..191,350..>356)
    /gene="cut-1"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CA10075.1"
    /db_xref="GI:3970656"
    /db_xref="SPTREMBL:O96776"
    /translation="PIQFATIGDPVYHKWTCDESEVDFCAVHSCFVDDGNDFVEI
    LNDGCGALDKYLLNLEYPFDL"
    1..191
    /gene="cut-1"
    /number=1
    192..349
    /gene="cut-1"
    /number=1
    350..356
    /gene="cut-1"
    /number=2

Intron
    110 a 64 c 64 g 120 t

BASE COUNT
ORIGIN
    110 a 64 c 64 g 120 t

Query Match
    13.5%; Score 156.6; DB 3; Length 358;
    Best Local Similarity 87.7%; Pred. No. 3.6e-26;
    Matches 171; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

    496 CCAATTCATTTGCTATCATGTCGACCAAGTTATCATTAATGACATGCGATTCTGAA 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 1 CCAATTCATTTGCTATCATGTCGACCAAGTTATCATTAATGACATGCGATTCTGAA 60

    556 ACCGTTATACCTTTCGCGGTTGTCATTCCTGCTTGTGATGATGATGATGATGATGATG 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 61 ACCGTTATACCTTTCGCGGTTGTCATTCCTGCTTGTGATGATGATGATGATGATGATG 120

    616 ACTGTGAAATTCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 121 ACGGTGAAATTCATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 180

    676 GAATATCCAACAGAT 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    Db 181 GAATATCCAACAGAT 195

RESULT 19
CEP53F1 39478 bp DNA linear INV 11-DEC-2001
LOCUS CEP53F1
DEFINITION Caenorhabditis elegans cosmid F53F1, complete sequence.
ACCESSION Z81088.1 GI:1627965
VERSION Z81088.1
KEYWORDS HTG.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

```









Db	39377	CACCTTAACCTTACGCAAAAGCTGAGACATGCTATTCACGATATGGGAATG-----	39327
OY	555	AACCGTGAATCTTCTGCGCGGTTGCCATTCCTGCTTGTGATGATGTAACGGTGA	614
Db	39336	TACACACACTGACATGAGGAATGCTGTCAAAAAGTGTTCTGTACCGATGGCGATGTA	39267
OY	615	TACCTGGAAATTCCTAAATGCGTAGGATGATGCTCTTATTAATATTTGCTTAATATTT	674
Db	39266	AGATATATCTGTGGTTGATTTGATGATGATGTCACACTATCCGTTTGTCTAGCGACT	39207
OY	675	GGAAATTCACACAGATTTAAATGCTGGCCAGGAAGAC---TCACGTATCAAAATATGCGGA	731
Db	39206	CTCTTATGATGCGTCTTTGATGATGAGCAGCATTCATCAGTACGATCGTCTTCAAGTATGCTGA	39147
OY	732	TCGATCAGACGCTTTTCTATCAATGCCAGATCAGTAT	767
Db	39146	CTCCAAATCAGCTATATCTTACATGTCAAATTCGTTT	39111
RESULT 21			
LOCUS	CEE04D5	31536 bp	DNA linear INV 23-JAN-2002
DEFINITION	Caenorhabditis elegans cosmid E04D5, complete sequence.		
ACCESSION	266496		
VERSION	266496.1	GI:1041307	
KEYWORDS	HTG; Cuticlin-like protein.		
SOURCE	Caenorhabditis elegans.		
ORGANISM	Caenorhabditis elegans. Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.		
REFERENCE	1	none.	
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED	99069613		
PUBMED	9851916		
REMARK			
REFERENCE	2	The C.elegans Sequencing Consortium.	
AUTHORS		2 (bases 1 to 31536)	
TITLE		McMurray,A.A.	
JOURNAL		Direct Submission	
COMMENT		Submitted (25-OCT-1995) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jesus@sanger.ac.uk or rwenematode.wustl.edu	
		Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.	
		Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.	
		IMPORTANT: This sequence is not the entire insert of clone E04D5. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.	
		The true left end of clone E04D5 is at 1 in this sequence. The true right end of clone E04D5 is at 2110 in sequence 248585.	
		The true left end of clone E04D5 is at 1 in this sequence. The true right end of clone E04D5 is at 31437 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence 249070.	
		The end of this sequence (31437..31536) overlaps with the start of sequence 248585.	
		For a graphical representation of this sequence and its analysis see: - <a href="http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=E04D5">http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=E04D5</a>	
		IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.	

Query Match

7.2%; Score 83.8; DB 3; Length 31536;

	Best Local Similarity	56.1%	Pred No.	4e-09:						
	Matches	179;	Conservative	0;	Mismatches	137;	Indels	3;	Gaps	1;
QY	480	TGGACCAACCGGTACCAGTTCATTTGCTATCATTTGGTCAGCAGTTATTCATAATG	539							
Db	27783	TGAAGCCGGTGATCACACTTAATAGTAGACGCTCGAATGCAGATCAGATCATCAAAATG	27842							
QY	540	GACATGCGATTCTGAAACCGCTTGATCTCTTGTGGCGGTTTCCATTCCTGCTTTGTCGA	599							
Db	27843	GACTTGCGCTTCCTGAACCTTGAGAAATGTTACTGTATGTAAGAACTTCATCATCTACGATTGA	27902							
QY	600	TGATGTGAACGGTGTACTGTGAAATTCTAAA---ATGCTGATGGAATGGTCTTGATAA	656							
Db	27903	TGAATGACACAAGGTGGACCAACCAGTAACGTGTATTTGATGCCAATGATGTCTTGATGATGG	27962							
QY	657	AATATTTGCTAAATATTTTGAATATTCACAGATTTATGCTGGCCAGAAGCTCAAGT	716							
Db	27963	GGTGATCTCTAGAAATCTTGATAGTACACTCTGACTTGACAGCTGGAAAACTGCAGCAGT	28022							
QY	717	ATACAAATATGCGGATGATCACAGCTTTTCTATCAATGCAGATCAGTATTAACATTAA	776							
Db	28023	CTTCAAAATTTCTGTGCAACAAGCTGGCTTTATTTCAATTTGTCAAATTAACAATCAACAA	28082							
QY	777	AGAACCAATATGCGCATGT 795								
Db	28083	AGACGTAATATPATGATGT 28101								
RESULT 22										
CDF53B6/C										
LOCUS			32412 bp	DNA	linear	INV 26-APR-2002				
DEFINITION			Caenorhabditis elegans cosmid F53B6, complete sequence.							
ACCESSION			281086							
VERSION			281086.1	GI:1627952						
KEYWORDS			HMG; Cuticulin; Human platelet tetraspan antigen like; Initiation factor associated protein; Thrombospondin like.							
SOURCE			Caenorhabditis elegans.							
ORGANISM			Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.							
REFERENCE			1							
AUTHORS			none.							
TITLE			Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium							
JOURNAL			Science 282 (5396), 2012-2018 (1998)							
MEDLINE			99069613							
PUBMED			9851916							
REMARK			The C.elegans Sequencing Consortium.							
REFERENCE			2 (bases 1 to 32412)							
AUTHORS			White,S.							
TITLE			Direct Submission							
JOURNALS			Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or rw@nematoide.wustl.edu							
COMMENT			Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information. Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. <b>IMPORTANT:</b> This sequence is not the entire insert of clone F53B6. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone F53B6 is at 1 in this sequence. The true right end of clone F53B6 is at 5116 in sequence Z81523. The true left end of clone F32H2 is at 32306 in this sequence. The true right end of clone T08G11 is at 7984 in this sequence. The start of this sequence (1..105) overlaps with the end of sequence							

280220.  
The end of this sequence (33306..32412) overlaps with the start of  
sequence 281523.  
For a graphical representation of this sequence and its analysis  
see: [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=F5386)  
[name=F5386](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=F5386)  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
the specified clone. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring submissions.  
Location/Qualifiers

## FEATURES

## source

1..32412  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="I"  
/clone="F5386"  
join(1863..1946,2437..2628,2969..3154,3205..3308,  
3430..3589,3804..3854)  
/gene="F5386.1"  
join(1863..1946,2437..2628,2969..3154,3205..3308,  
3430..3589,3804..3854)  
/gene="F5386.1"  
/note="predicted using GeneFinder  
Similarity to Human platelet-endothelial tetraspan antigen  
(SMLEPT3\_HUMAN), contains similarity to Pfam domain:  
PF00335 (Transmembrane 4 family), Score=30.0,  
E-value=5.8e-09, N=1"  
/codon\_start=1  
/protein\_id="CAB03120.2"  
/db\_xref="GI:11065632"  
/db\_xref="SPTREMBL:O9XVW9"  
/translation="MGALDSAYGANGRIKSYVTALISLIFSCICYIMILAR  
KQYALVPSLYVDGRLVITISLILNYLICFALIEKRCFVSCAVASIVAV  
MLIGGICLGNRQD/LHYTPNLKMLTSLRELYGHDKGTESDALQSNFKCG  
NCTDNAQIKWTKSWHORAPKLIPESSCISEIEIRCSNPFDDAPPPYTSICE  
PLONDLVHNAVSMCLTNATVQIIPSVAGCMYSKLIRK"  
complement(join(4270..4339,4493..4547,5071..5395,  
5368..5662,5891..6148,6204..6752,6818..7287,7546..7676,  
8104..8184,12576..12775,13246..13384,13931..14132,  
14183..14294,14341..14457,14994..15155,15198..15311))  
/gene="F5386.2"  
complement(join(4270..4339,4493..4547,5071..5395,  
5368..5662,5891..6148,6204..6752,6818..7287,7546..7676,  
8104..8184,12576..12775,13246..13384,13931..14132,  
14183..14294,14341..14457,14994..15155,15198..15311))  
/gene="F5386.2"  
/note="predicted using GeneFinder  
contains similarity to Pfam domain: PF00047  
(Immunoglobulin domain), Score=24.4, E-value=1.7e-06, N=1  
CDNA EST EMBL:214404 comes from this gene  
CDNA EST yk153a1.5 comes from this gene  
CDNA EST yk195c9.3 comes from this gene  
CDNA EST yk195c9.5 comes from this gene  
CDNA EST yk286g7.3 comes from this gene  
CDNA EST yk286g7.5 comes from this gene  
CDNA EST yk323e3.5 comes from this gene  
CDNA EST yk353h1.3 comes from this gene"  
/codon\_start=1  
/protein\_id="CAB03121.1"  
/db\_xref="GI:4008366"  
/db\_xref="SPTREMBL:P90884"  
/translation="WKSTVYVFLPYLLISFHVDAISMAAMSPWSSCTTGGGVS  
ROLRCLTKSGSGEVRFRKVCQAKTCEKSKSLARDTICGGEIVSRGCEVYGRSLT  
GANFLMYDDGTPCOAATSRVCSKSGCOIYCDGLISSFFRFDATGGGGGDCDN  
GKFIWKEEYATACASNCDDIVDMAGRSRTASTOPTVCYNATLGRVVEPKICADK  
LAPKVEARPCPMILICPSRMEADKSKSTASGQGVRRRHVACVGGSDCGGRRRQ  
ENGTRPAEBNCVSTSGRWEAKSKSTASGQGVRRRHVACVGGSDCGGRRRQ  
TTCTAGIPCSTATNSLDWNDRAVLDGNFSGMDNMDQAPRLVAGESTSGTG  
VMSRTVECAVAVPISAPILKIPMSECOQEQPKLESCVRSCLQEDSKISEDAFY  
OMRGDTMVCASVSLGCGKOKALKICIOVSTGKVSQSCDARRRPEKSRPCNOHPC  
PFWLTKSKSDCSMGSGCTARRSVCAQTVSKTGDADHAIYLRDRCKFRKPOEETC  
NVVACPATVWSLNRKRNKILINKLTKIQWECSSCSGSRRRQVWCIEIDSRKTO  
RRPVEDCANTRKPTQIVEVCSFGSCSPRELLSNRFEDONAEOKKILIGIYVATLYOGT

gene

CDS

gene

CDS

gene

CDS

gene

SIKINCPARKFDKKIYWKNGKRIKINDAHIKVANSGLNLRVFNARMEDAGVCEFTDR  
LOGNYLNRKYEDPASPASRYDLAKPOIPSTKNROROVSKEDVLEQASVILKMYVSL  
IFALLTAPNDEKAREOLRKYGNELVARMOIGMHSROKTCVAGVOAQASCKYTFH  
GFRINDVNSICISLASVRRPEPRRREDCPPMPSQMSQSECSQVCSMLAKORNY  
NRRFNSTGVDDOHCDDITRRPATTTNDPCPNQKAKRRKRSIDMSGCSSECTGGVQURL  
SCVWISSGRPACRNCBOMRRPHASRACVADEFLPPCMPASALYORDASQDQSRFCD  
IKLPHSCDLSLEVRQKCCSTCFVEFRKK"  
complement(join(18772..18944,19232..19286))  
/gene="F5386.9"  
complement(join(18772..18944,19232..19286))  
/gene="F5386.9"  
/note="CDNA EST EMBL:T02267 comes from this gene"  
/codon\_start=1  
/protein\_id="CAD30440.1"  
/db\_xref="GI:20338942"  
/translation="MANOGRFLVFFVVLLEFATVTAHHSHSHSHSHHHHNGYTG  
GGYGVQPRYGGGGYNGYNGYKRYOYNGKRRK"  
complement(join(13992..20022,20241..20283))  
/gene="F5386.8"  
complement(join(13992..20022,20241..20283))  
/gene="F5386.8"  
/note="CDNA EST yk553f11.3 comes from this gene  
CDNA EST yk553f11.5 comes from this gene"  
/codon\_start=1  
/protein\_id="CAB54255.1"  
/db\_xref="GI:5824510"  
/db\_xref="SPTREMBL:O9U3E7"  
/translation="MDFGILLVFLMAVAVGTGAGISVSFSSHSHEPRHLLGGCGFPNY  
GGYQSNYWGKRR"  
complement(join(20516..20602,20648..20953,20999..21424,  
21609..21818))  
/gene="F5386.4"  
complement(join(20516..20602,20648..20953,20999..21424,  
21609..21818))  
/gene="F5386.4"  
/note="predicted using GeneFinder  
contains similarity to Pfam domain: PF00635 (MSP (Major  
sperm protein) domain), Score=74.1, E-value=9.1e-19, N=1  
CDNA EST yk193h3.3 comes from this gene  
CDNA EST yk411f5.3 comes from this gene  
CDNA EST yk193h3.5 comes from this gene  
CDNA EST yk411f5.5 comes from this gene  
CDNA EST EMBL:AW057173 comes from this gene  
CDNA EST EMBL:AW057232 comes from this gene  
CDNA EST EMBL:AW057344 comes from this gene  
CDNA EST EMBL:AW057361 comes from this gene"  
/codon\_start=1  
/protein\_id="CAB03117.1"  
/db\_xref="GI:3877449"  
/db\_xref="SPTREMBL:O9XVNI"  
/translation="MIRSRGQPFRRSKCDQSDINRKNQFVLEKEARYCASIARTML  
VPLAIVTSLPVAALIAIFCAKRNRTYVAKRNNKSSKSAKSKSTPGASIKGSRMR  
SKAKHVSRRSSKSKGKSGKSGKSSKSKSKSKKSKKSKKSKKSKKSKKSKKSKKSKK  
TGVSDKQNSKSSKSSKSSKSKSKRNRLDSDAQKMKSGKSGVALIPKTOQTGS  
OVGSHLAEVNSIKHSKEMNAPAKLQYQTLGQVNOLELKTMSERKAYKIKGSDNSL  
YRVNVEGFAPRRSVKIDVLRNGEOKTDLVLTANANGSTNPHEAFANQAREHM  
MVEPLVAA"  
23555..24160  
/gene="F5386.5"  
23555..24160  
/gene="F5386.5"  
/note="Similarity to Rat Initiation factor 2 associated  
glycoprotein (SW:12A6\_RAT)"  
/codon\_start=1  
/protein\_id="CAB03119.1"  
/db\_xref="GI:3877451"  
/db\_xref="SPTREMBL:P90883"  
/translation="MILSDQNFLOTQMKPEPQATQSKTSECEFGNSNEYKPGSLN  
GQSLAQRIRHEKVEPIYKGGEOARMENETVATFEFGSGKGYFHDMDTSHYMKRP  
ELABEKIPLRQKSKGLNLDRFATLARCVIDLNGETKIKYIMALKGKCSYSTAAR  
LMLTLTPELDRQKMMVACIIQSFAGITIGRAVLLNLII"  
join(26258..26342,26501..26802,26846..26926,26974..27107,  
27153..27288,27334..27459)







The end of this sequence (36203..36306) overlaps with the start of sequence 299276.  
For a graphical representation of this sequence and its analysis see: <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=M142>  
IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

## FEATURES

```
source
1..36306
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="III"
/clone="M142"
join(3511..3570,4639..4750,4821..5201,5689..5795)
/gene="M142.1"
join(3511..3570,4639..4750,4821..5201,5689..5795)
/note="cDNA EST Yk258a1.3 comes from this gene"
cdna EST Yk258a1.5 comes from this gene
cdna EST Yk352a11.3 comes from this gene
cdna EST Yk352a11.5 comes from this gene"
/codon_start=1
/protein_id="CA97807.1"
/db_xref="GI:3878736"
/db_xref="SWISS-PROT:O10658"
/translation="MKABDQOQSIAPGATFPSPQMRPPPTQATITTEALLAKNOI
TPNDVALPEITGGLCSPSANVNIETFEQJRLDTEVLEIAKPEETENIQ
QAESARYVRKAPNFKLTGATVEFGEDVPIHFEMIEHFFKDRLLKCFDEF
GFCMPSRNCEIYEFPOLSQOLMDMINNPNETRSDSYFEVKNKLVMAKADYSYD
A"
join(8540..8572,9335..9416,9468..9679,10373..10721,
10772..10974,11033..11175,11871..12501)
/gene="M142.2"
join(8540..8572,9335..9416,9468..9679,10373..10721,
10772..10974,11033..11175,11871..12501)
/note="M142.2"
/note="Predicted using GeneFinder
contains similarity to Pfam domain: PF00092 (von
Willebrand factor type A domain), Score=155.0,
E-value=4.3e-43, N=1
cdna EST Yk518d1.5 comes from this gene"
/codon_start=1
/protein_id="CA97806.1"
/db_xref="GI:3878735"
/db_xref="SPTREMBL:O21540"
/translation="MRPIPYDISITSFSLSLILCSANFIDNGLVDSLEHCVTHK
AVEVILLDSGSGIDDTFKQLSPAMHLASRLNISGDSHMLIOYAEPTKLEFSIG
QFNHPQLEMAIORIEYSGATNTGOALRLTLERGLGAGRGIPKVAIVITDGSODD
VSEPSQLRDADMYAIGYTNLVNHOHQMGNPYRVFTYESFQDLDAALDSLTM
SKCTERRETPETIGCPDRIGVAKSTKQPEEGNVFMDHTHDEECAGKEKPPDSKS
IGLVPSACNVHRVSLNPKGIFVEASTVFNHSLFMTDTQTVKQCFEMKADKRS
TVPLSVMTTVREQIYQMPQCAVILRKAGPDGPIVREFATLCESYVHRECELEVA
KDTFGMLHSCYVNGDYGDRVLDLDSNGCDLAVLSTPDYDLSLRLAKPVHVEY
ADRPVLOFOCOLITLKYDGCSEGITPPQCKKLKPEDGHHNHHEKRRKRLVRBLAD
GVGTIDVTSVTVLEOERACQOVTTYESPSTSTNOCKLENS"
join(17427..17468,18628..18766,20758..20915,22487..22677,
22840..23053)
/gene="M142.4"
join(17427..17468,18628..18766,20758..20915,22487..22677,
22840..23053)
/note="VAB-7 even-skipped homologue, contains similarity
to Pfam domain: PF00046 (Homeobox domain), Score=111.1,
E-value=7.1e-30, N=1"
/codon_start=1
/protein_id="CA97809.1"
/db_xref="GI:3878738"
/db_xref="SWISS-PROT:Q93899"
/translation="MQSFIDSLIGVAKPSLVEMVAASRASFSPPEQOHHDPKGV
VAAAAAAGRRHHPYDNRDQGMRRRTAFSREOIGRLREFEAKENYVSRKTRGEIA
AEINLPEGITIKVWFQNRKMKDKRQVGGIAMPPPQMAAYMLNPFVEMMKTTAASQ
```

```
gene
join(28250..28366,28422..28499,28558..28991,30672..30966,
31500..31763,32933..33043)
/gene="M142.5"
join(28250..28366,28422..28499,28558..28991,30672..30966,
31500..31763,32933..33043)
/note="Predicted using GeneFinder
cdna EST EMBL:T01774 comes from this gene
cdna EST Yk470a9.3 comes from this gene
cdna EST Yk470a9.5 comes from this gene
cdna EST Yk476b6.5 comes from this gene
cdna EST Yk563c12.3 comes from this gene
cdna EST Yk662c9.3 comes from this gene
cdna EST Yk679e8.3 comes from this gene
cdna EST Yk563c12.5 comes from this gene
cdna EST Yk584c4.5 comes from this gene
cdna EST Yk662c9.5 comes from this gene
cdna EST Yk679e8.5 comes from this gene"
/codon_start=1
/protein_id="CA97808.1"
/db_xref="GI:3878737"
/db_xref="SPTREMBL:O21541"
/translation="MASCNLTASNAVAGCEVETITNGLSARGVTSFDTTRKVLVD
TKEMAIKPKPIRIFNSHLKNTIVLSALESEKSPKAKTEQDFRANPNVGTETTERL
DKTGLGELPLMKSPATISIGQAVYIOLKRTIPDTCGCDIVGLIVLHKRYEVD
VTKDAKATGDESRAESAVLQOVKILSKPTEYTAARPILDTFGALVANSSEMSDT
DDIIPOLASDTLALSMFOAQOEKIDLOGLIEKIDLOQFOWDDTSKRLVAE
GVAALGESBARARIGCVSPPLVGFHEHEKTEQIDOLTFFEPDRGLGKLEPFEF
VHYDKRHPDLPAELAKFEVITLADPEFLACIKASHIRLLGKSDVVLCTGTAI
MEDYASRLMAMHSTFEPFRANNLANDSECFANVOTLTFC"
join(34039..34101,34448..34750,35530..36306,
299276.1:105..698,299276.1:1364..1566,299276.1:1666..1772,
299276.1:2495..2795,299276.1:3714..3928,
299276.1:4056..4104,299276.1:5846..5958,
299276.1:6478..6729,299276.1:7169..7338)
/gene="M142.6"
join(34039..34101,34448..34750,35530..36306,
299276.1:105..698,299276.1:1364..1566,299276.1:1666..1772,
299276.1:2495..2795,299276.1:3714..3928,
299276.1:4056..4104,299276.1:5846..5958,
299276.1:6478..6729,299276.1:7169..7338)
/note="Similar to Zinc finger, C3HC4 type (RING finger)
cdna EST Yk49h8.5 comes from this gene"
/codon_start=1
/protein_id="CA97810.1"
/db_xref="GI:3878739"
/db_xref="SPTREMBL:O45962"
/translation="MAPTGGGQGOEYVLCGSCICNHFNEFLFVSLICGIVTRKCAE
KPNQTPKCHMDMKTTTHSPSEYPMNVALLSVIFPRQCTLTSGAVSEAKRDQSLI
QIAKFEHEADSEKGVSSREISRTLOKRLKALQOMRVDRDLKTLKRCISEVY
MIIILISQENTVSSOLMAGVARGQGLPMQDVLILMLTLETGTCIRKINLY
MYVQVITLASYPOVSKTCVGVVOLLYADSCFNVLRKDGSSLMOLKEEPTASIR
EHSOVLIVAFESGLRIGRQVWGAALLYADOSHMSQSIDTKLOSKNSOOGVEELRA
LASQTSMLYPAIRYPIYQVIFCLPEFAGIEHEDTSRMIGDALHQRILILKHCSD
DLKMKPEEKRGVILDAEVPFGMGGGGGGAGAGTIGLHLPLISQIDTGRSISIT
NPKDNSHNSPQTPPKQPRKRYOMGIPRPMGSSAPPEIPSHQOQPPQPFNSQHL
PQREFRGGRGAGAPPPEPPQPMPLIGIDMGPAGMPOKTEVLITADGQVNSTPORVIM
QSPVHLPGGVWMIPOQVMPPOSMTVGGPAGPMQKPTSPVQVPRMTWTASP
TGSVIYPAASPCQPPHTWIOSIGVKKRSNPLKTYKISFLNFPDPLILRLTKK
EKGADIDLEFEKIKSTDFFKYPSSFSSTQDNMPHMDRGSGAGYWGFGYLRSGADA
EQLAKYTEILKRLQDSDDDDPDGIGHVSTVASVSLDDRDHHPILMTIPVPTID
LPAIFLISFANPTEETMTMIGVAMQNPRAPSLAPSSNQPMNVNASSATVQAEEN
RKILIPFLKRYKMTLEKYSTCFHYTLKLDYVYFVYLANLNASRPPRRRAATIPQ
VPLPMPVOVPOVIVPAENPNPNVPPPPPPPOCPMLVDSAGILTPIRILVAHON
VNSNLSLDKIVDKVIERISQAGNASSPNAHLRKLRLRAESQMAHDDPYTNNCLLAL
QOVDMELOOLHLPYVEG"
```

BASE COUNT	ORIGIN	Query Match
11616 a	6626 c	6322 g 11742 t

5.4% ; Score 62.6; DB 3; Length 36306;





[illegible]

once, or longer because we arrange for a small overlap between neighbouring submissions.  
The true left end of clone W01A8 is at 1 in this sequence. The true right end of clone W01A8 is at 11258 in sequence 271262.  
The true left end of clone F2206 is at 25181 in this sequence. The true right end of clone F15c11 is at 12341 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence 271260.  
The end of this sequence (25181..25284) overlaps with the start of sequence 271262.  
For a graphical representation of this sequence and its analysis see: <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=W01A8>  
IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

## FEATURES

Location/Qualifiers

1..25284

/organism="Caenorhabditis elegans"

/db\_xref="taxon:6239"

/chromosome="1"

/clone="W01A8"

complement(join(5595..5813,6012..6100,6360..6495,7471..7530,7729..7856,7934..8072,8122..8142))

/gene="W01A8.2"

complement(join(5595..5813,6012..6100,6360..6495,7471..7530,7729..7856,7934..8072,8122..8142))

/gene="W01A8.2"

/note="predicted using GeneFinder"

CDNA EST yk275h2.3 comes from this gene

CDNA EST yk309g11.3 comes from this gene

CDNA EST yk275h2.5 comes from this gene

CDNA EST yk309g11.5 comes from this gene"

/codon\_start=1

/protein\_id="CA95853.1"

/db\_xref="GI:3880399"

/db\_xref="SPTREMBL:Q23094"

/translation="MKPKKRIANKCVPTGEIVDKAQSALFSDTERGLHTIRKPGAKSCVAIGDSEVDVAIGAAPREGANELLISYMSALGRKNELOFGAKSRSKYVILDTKRTIDETISVSAKDPDEICEKVEYEDKFCGPGYPLHGLRGLCFNNVCTDLKFTPGKRGKRYGCGSTNCLVKATIVNISESTICSTIOEAFKSHVDCYINCFCFVCKTEKLAFLGSYDWTDFLSYAAVKQVLAIIKRCGIFSCIFKEDY"

join(9030..9125,9180..9776,9822..9899,9955..10158,10206..10307,10359..10418,11007..11117)

/gene="W01A8.1"

join(9030..9125,9180..9776,9822..9899,9955..10158,10206..10307,10359..10418,11007..11117)

/gene="W01A8.1"

/note="predicted using GeneFinder"

CDNA EST yk75g3.3 comes from this gene; CDNA EST yk75g3.5 comes from this gene

CDNA EST yk163b12.5 comes from this gene; CDNA EST yk442h5.3 comes from this gene

CDNA EST yk442h5.5 comes from this gene; CDNA EST yk303h10.3 comes from this gene

CDNA EST yk303h10.5 comes from this gene; CDNA EST yk206b12.5 comes from this gene

CDNA EST yk239h1.5 comes from this gene; CDNA EST yk281d10.5 comes from this gene

CDNA EST yk302e2.5 comes from this gene; CDNA EST yk303e9.5 comes from this gene

CDNA EST yk339h11.5 comes from this gene; CDNA EST yk365h6.5 comes from this gene

CDNA EST yk383d9.5 comes from this gene; CDNA EST yk417d8.5 comes from this gene

CDNA EST yk428d4.5 comes from this gene; CDNA EST yk460c8.5 comes from this gene

CDNA EST yk477f1.5 comes from this gene; CDNA EST yk482d12.5 comes from this gene

CDNA EST yk525f10.3 comes from this gene; CDNA EST yk340b10.3 comes from this gene

gene

CDS

CDNA EST yk679g8.3 comes from this gene; CDNA EST yk239h1.3 comes from this gene  
CDNA EST yk281d10.3 comes from this gene; CDNA EST yk302e2.3 comes from this gene  
CDNA EST yk303e9.3 comes from this gene; CDNA EST yk339h11.3 comes from this gene  
CDNA EST yk383d9.3 comes from this gene; CDNA EST yk417d8.3 comes from this gene  
CDNA EST yk428d4.3 comes from this gene; CDNA EST yk460c8.3 comes from this gene  
CDNA EST yk477f1.5 comes from this gene; CDNA EST yk482d12.5 comes from this gene  
CDNA EST yk525f10.3 comes from this gene; CDNA EST yk340b10.3 comes from this gene

gene

CDS

CDNA EST yk679g8.3 comes from this gene; CDNA EST yk239h1.3 comes from this gene  
CDNA EST yk281d10.3 comes from this gene; CDNA EST yk302e2.3 comes from this gene  
CDNA EST yk303e9.3 comes from this gene; CDNA EST yk339h11.3 comes from this gene  
CDNA EST yk383d9.3 comes from this gene; CDNA EST yk417d8.3 comes from this gene  
CDNA EST yk428d4.3 comes from this gene; CDNA EST yk477f1.3 comes from this gene  
CDNA EST yk482d12.3 comes from this gene; CDNA EST yk525f10.3 comes from this gene  
CDNA EST yk540b10.5 comes from this gene; CDNA EST yk580a3.3 comes from this gene  
CDNA EST yk580a3.5 comes from this gene; CDNA EST yk612a9.5 comes from this gene  
CDNA EST yk612e9.5 comes from this gene; CDNA EST yk615c12.3 comes from this gene  
CDNA EST yk615c12.5 comes from this gene; CDNA EST yk643h10.3 comes from this gene  
CDNA EST yk643h10.5 comes from this gene; CDNA EST yk670g10.5 comes from this gene  
CDNA EST yk679g8.5 comes from this gene"

/codon\_start=1  
/protein\_id="CA95854.1"  
/db\_xref="GI:3880400"  
/db\_xref="SPTREMBL:Q23095"  
/translation="MTDYEQPSVSDQAOASQYDVYDVGNAVYQTAIAYTKTEPHPLNLTLSAEKSVTVGNIAOKAYIDTSTNYKPKRTATENAISYGERKATVESCROMALVGGTFGIGAANVLTFSLASAGALVLEQVDSARKLSSALSTKELELVGHRIFSAHQOIRAMVPEVERKINTENTNLILDLDAVQKINLIEVPPVNLTIQORVNIASLIQVGSNKADHVIDPENERARNYLDLSQSFVLLDIVEKKTWVLEKSNELSTVDFPKTLEEEAOKYKVAPEEMKIKHIOSELSLOSLRKGONVGGDKIDSTVDLEMLKKKNTDAEDYVYRDEVLMNEGQRIALSTWTSLLIISAENQFPEDLILIEELXPAPPPVPRNLRXRA"

join(18064..18085,18138..18325,18375..18520,18908..19295,19346..19435,19506..19721)  
/gene="W01A8.3"  
join(18064..18085,18138..18325,18375..18520,18908..19295,19346..19435,19506..19721)  
/gene="W01A8.3"  
/note="similarity to C.elegans cuticulin (SW:CU11\_CABEL)"  
/codon\_start=1  
/protein\_id="CA95850.2"  
/db\_xref="GI:15718223"  
/db\_xref="SPTREMBL:Q23097"  
/translation="RRIWRCSTPKVICAEKDIALDVTSKPFGNTFVKGRADKCKOSTANNKTSYSLPLKCGQMLRBSANPRGVNRYIVSFHFGFTTKNDRAFHVCFTVEPDEIVQNIIDVSMIPTELSDSVWMPKCESVARDPNPGPTLYAVNGDVFHVECTPADMGMLVKKCFVTDGDEHAAVDPDGCATDPFLSELSTYDASLMRAHASSO VFKYADSNQLFTFCQIRLCOKROMGCEVTPPCGVKKLSEASGDDGKRTREADRSD VEIVATSELLVLDPADRGCLAPSPFCVPRLLVPLPILITIVSLTVVSPALVIARQ NKKKELDIMOSGRY"

complement(20058..20321)  
/gene="W01A8.4"  
complement(20058..20321)  
/gene="W01A8.4"  
/note="CDNA EST yk100d3.3 comes from this gene  
CDNA EST yk100d3.5 comes from this gene  
CDNA EST yk607d1.3 comes from this gene  
CDNA EST yk607d1.5 comes from this gene"

/codon\_start=1  
/protein\_id="CA95851.1"  
/db\_xref="GI:3880397"  
/db\_xref="SPTREMBL:Q23098"  
/translation="MDPAMFRWISADMTQAEFFFPRTYVFLYGVPAFYITRLM FVPMDKSEACLDLCKLIMWDRQRTLOTGGSGLFAPTIGIDF"  
join(21264..21365,21414..21568,21618..21843,21891..22068,22134..22408)  
/gene="W01A8.5"  
join(21264..21365,21414..21568,21618..21843,21891..22068,22134..22408)  
/gene="W01A8.5"

gene

CDS

gene

CDS

/note="predicted using GeneFinder  
 cDNA EST yk67a6.3 comes from this gene  
 cDNA EST yk67a6.5 comes from this gene  
 cDNA EST yk604e4.3 comes from this gene  
 cDNA EST yk604e4.5 comes from this gene"  
 /codon\_start=1  
 /protein\_id="CA95849.1"  
 /db\_xref="GI:3880395"  
 /translation="METDGDPTNEIFGSSMAFVANHVTPEERSPLDRIAEGRNG  
 GWSDDLSLDGVKIKYKTKETKATVIRNIVKKSIEIKRIDETIKIEHKEI  
 ILPECKNWKIEAGHHNITPPPCQVDPNDWSSIMGKIVDHHGSIINTLPSLRPT  
 ETLFENIALESKTHENAVYTKLLTAATPDDSDNIPVPIYRMRKACATLEBMPALN  
 ELEAAVSVKVDITETETPTDEKSIAGIIFSECOCNRYRRCODYPNVAAVOL  
 FVDPLRTREHAIPEVASEVELINFTKDNOPSTSSN"  
 complement(join(22578..22756,22809..22983,23343..23481,  
 23564..23757,24032..24154,24283..24565,25191..25219))  
 /gene="W01A8.6"  
 complement(join(22578..22756,22809..22983,23343..23481,  
 23564..23757,24032..24154,24283..24565,25191..25219))

Query Match 4.8%; Score 56.2; DB 3; Length 25284;  
 Best Local Similarity 51.2%; Pred. No. 0.012;  
 Matches 189; Conservative 0; Mismatches 168; Indels 12; Gaps 2;

OY 415 ATGACACGCTTTTCAACTCAATATGCGCCAGTACGCTTATGAAATTTTG 474  
 DB 18912 ATCCCAACGACTGACTATGTGACAGTATGTAATGCCAAATGCGATTTTCAGTTCGT 18971  
 OY 475 GATGATGAGCAACCGGTCAACAGGTTCAATTTGCTATCATTTGTCAGCCAGTTTATCAT 534  
 DB 18972 CGTATGATGCTCAATATGCTCAACGTTTACATATGCCAATGTCGAGATATTGATTTTCAT 19031  
 OY 535 AAATGACATGCAATTCGTGAACCGTTGATACCTTTCTGCCGGTGTTCATTCCTGCTT 594  
 DB 19032 GTATGGGAAATGT-----ACACCGGCTGATATGGAATGCTTGCATAAAGTCTTTT 19082  
 OY 595 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 654  
 DB 19083 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 19142  
 OY 655 AAATATTTGCTAATATATTTG---GATATTCACACAGATTTAAATGCTGCGCCAGAAAGCT 711  
 DB 19143 CCATTTCTTCTCAGTGCAGTGTCTTACAGATGCTTCACTTAATGGAGCTCATGCATCAAGT 19202  
 OY 712 CAGCTATACAAATATGCGGATCGATCAGCTTTTCTATCAATGCCAGATCATATTACC 771  
 DB 19203 CAGATTTTCAAAATGACGACATCAATCAATTAATTTCACTTGTGATGATTAAGTGTGT 19262  
 OY 772 ATTAAGAA 780  
 DB 19263 CAGAAACAA 19271

RESULT 27  
 AC084450 59533 bp DNA linear INV 04-NOV-2000  
 LOCUS Caenorhabditis briggsae cosmid CB022A21, complete sequence.  
 DEFINITION AC084450  
 AC084450  
 VERSION AC084450.1 GI:11094900  
 KEYWORDS HTG.  
 SOURCE Caenorhabditis briggsae.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 59533)  
 AUTHORS The C. briggsae Genome Sequencing Center.  
 JOURNAL Washington University Genome Sequencing Project  
 REFERENCE 2 (bases 1 to 59533)  
 AUTHORS Waterston,R.  
 TITLE Submitted (04-NOV-2000) Department of Genetics, Washington  
 JOURNAL

## COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 Submitted by:  
 Genome Sequencing Center  
 Department of Genetics, Washington University,  
 St. Louis, MO 63110, USA  
 e-mail: jsplethwatson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we provide a small overlap between  
 neighboring submissions.

FEATURES  
 source location/Qualifiers  
 1..59533  
 /organism="Caenorhabditis briggsae"  
 /strain="GuJarat G16"  
 /db\_xref="taxon:6238"  
 /clone="CB022A21"

BASE COUNT 18802 a 10893 c 11126 g 18712 t  
 ORIGIN

Query Match 4.8%; Score 55.4; DB 3; Length 59533;  
 Best Local Similarity 62.8%; Pred. No. 0.018;  
 Matches 86; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 271 AATCCAGTGTGATTTTGTAAACAACAGTGTGCTTTCATTTGTTATTTGTT 330  
 DB 5291 AACCCAAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52850  
 OY 331 ACCAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390  
 DB 52851 ACCAAATGACAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52710  
 OY 391 AGTGCACAGATGAGGT 407  
 DB 52711 ACCGACGAATTAATGT 52727

RESULT 28  
 U41264/c 28396 bp DNA linear INV 22-MAY-2002  
 LOCUS Caenorhabditis elegans cosmid F10E7, complete sequence.  
 DEFINITION U41264  
 U41264  
 ACCESSION U41264.1 GI:1086829  
 VERSION  
 KEYWORDS HTG.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 28396)  
 AUTHORS Waterston,R.  
 TITLE Genome sequence of the nematode C. elegans: a platform for  
 JOURNAL Investigating Biology. The C. elegans Sequencing Consortium  
 MEDLINE Science 282 (5396), 2012-2018 (1998)  
 PUBMED 9851916  
 REFERENCE 2 (bases 1 to 28396)  
 AUTHORS Pauley,A.  
 TITLE The sequence of C. elegans cosmid F10E7  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 28396)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-NOV-1995) Robert Waterston  
 REFERENCE 4 (bases 1 to 28396)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-2001) Department of Genetics, Washington  
 REFERENCE 5 (bases 1 to 28396)  
 AUTHORS Louis, MO 63110, USA  
 TITLE Direct Submission

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

JOURNAL

## COMMENT

Submitted (27-SEP-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
6 (bases 1 to 28396)  
Waterson, R.  
Direct Submission  
Submitted (22-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
Submitted by:

Genome Sequencing Center  
Department of Genetics, Washington University  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
email: rtw@emulab.wustl.edu and jessesanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

For a graphical representation of this cosmid sequence and its analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=F10E7;class=Sequence>

## NEIGHBORING COSMID INFORMATION

The 5' cosmid is T03G5, 200 bp overlap; the 3' cosmid is T07F8, 1000 bp overlap. Actual start of this cosmid is at base position 1 of F10E7; actual end is at 7817 of T07F8.

## NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneFINDER (P. Green and L. Hillier, personal communication), the large scale EST projects of Yui Kohara ([http://www.ddbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and the C. elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>), similarity to other proteins from Blastx analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans Genbank submissions, and personal communications with C. elegans researchers, TRNAS are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

## FEATURES

## source

1. 28396

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

/db\_xref="taxon:6239"

/chromosome="II"

/clone="F10E7"

complement(900..4554)

/gene="F10E7.8"

/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=F10E7.8;class=Sequence>"

## CDS

complement(join(900..1126,1176..1637,1685..1769,1985..3460,3510..3689,3736..3885,4119..4291,4518..4554))

/gene="F10E7.8"

/note="coded for by the following C. elegans cDNAs: YK4491.3, YK20F8.3, YK12B7.5, YK16912.3, YK16912.5, YK20F8.5, YK3696.3, YK3696.5, YK4491.5, YK361h1.5"

## gene

## CDS

/codon\_start=1  
/product="Hypothetical protein F10E7.8"  
/protein\_id="AA082421.1"  
/db\_xref="GI:1086830"  
/translation="MPAHLNWKACRITGDKERRSADDEHYEYSSSLAARSANVTK  
LKQLEAOPMTIPAGFTIHSSDNDKRDHQFTEEFESAVQKNTYIQDCKMES  
ADLEVRKLSARITILEFLOGSAVDFEGCLENEETFEYQDKRLMNPVIRTKRMDPES  
MYHVAIRNSFICYKNGVYQALCTLMLEIRTEFPSPNDIRSASSRSSRMSLAD  
LSDSEKRTROLPTVNDENMLRVILAGYHMHIVOLDRESGREDTDDLTETRAFRE  
EVOEPIENGNEPLILVFDMLHPPIYGAHPPIKTIKILITLLTGMEETVR  
KREKMTAGDIMEPTITVAASQAFKADDEHNRNMAHRGSPARGMAMARQAYND  
DSKDDAYSSESDSTASKKEDSPGASSTPSIPDTASNFRQSGEQPRGSPV  
FVYQKTLTPMRSKTNMADIEDFQKGRKTYNDFDGDGDTLLEGLPSCGAVKIL  
RNMTYSTLEIQVQKDEELNRYLSMKEDIPETKTELRLPNLSQICALAKVMY  
SSVPTKARHEGLANVLDCLPMEASDILNSISLSDSTSPSEGRALADINRK  
EIVKSLAIIILLVYKHKLNHIYOFEYICQOIYVNSIPLIKFLDONTKLIOSR  
EIVAYNPQCILYHYRNNEHPILODNIERRPRGASPYFMNRVYPAFNLILNK  
LVKANDVAKMLVFKSAPVLRKFKVNSVQLVLAIAIMQSRKGLQKRSKMDI  
ISATSRVHRMTDWMAFASIDIKRCDQKDSLIKASIERFHSRRSALYPOFAFRE  
NDAPWPGDDYLNRYDMRDFEVDICASHVLGANLIKGRHFKKDYEWLEQVFNASID  
MDKLIERGVEDLM"

## gene

## CDS

/codon\_start=1  
/product="Hypothetical protein F10E7.8"  
/protein\_id="AA082421.1"  
/db\_xref="GI:1086830"  
/translation="MPAHLNWKACRITGDKERRSADDEHYEYSSSLAARSANVTK  
LKQLEAOPMTIPAGFTIHSSDNDKRDHQFTEEFESAVQKNTYIQDCKMES  
ADLEVRKLSARITILEFLOGSAVDFEGCLENEETFEYQDKRLMNPVIRTKRMDPES  
MYHVAIRNSFICYKNGVYQALCTLMLEIRTEFPSPNDIRSASSRSSRMSLAD  
LSDSEKRTROLPTVNDENMLRVILAGYHMHIVOLDRESGREDTDDLTETRAFRE  
EVOEPIENGNEPLILVFDMLHPPIYGAHPPIKTIKILITLLTGMEETVR  
KREKMTAGDIMEPTITVAASQAFKADDEHNRNMAHRGSPARGMAMARQAYND  
DSKDDAYSSESDSTASKKEDSPGASSTPSIPDTASNFRQSGEQPRGSPV  
FVYQKTLTPMRSKTNMADIEDFQKGRKTYNDFDGDGDTLLEGLPSCGAVKIL  
RNMTYSTLEIQVQKDEELNRYLSMKEDIPETKTELRLPNLSQICALAKVMY  
SSVPTKARHEGLANVLDCLPMEASDILNSISLSDSTSPSEGRALADINRK  
EIVKSLAIIILLVYKHKLNHIYOFEYICQOIYVNSIPLIKFLDONTKLIOSR  
EIVAYNPQCILYHYRNNEHPILODNIERRPRGASPYFMNRVYPAFNLILNK  
LVKANDVAKMLVFKSAPVLRKFKVNSVQLVLAIAIMQSRKGLQKRSKMDI  
ISATSRVHRMTDWMAFASIDIKRCDQKDSLIKASIERFHSRRSALYPOFAFRE  
NDAPWPGDDYLNRYDMRDFEVDICASHVLGANLIKGRHFKKDYEWLEQVFNASID  
MDKLIERGVEDLM"

## gene

## CDS

/codon\_start=1  
/product="C. elegans RPL-33 protein (corresponding sequence F10E7.7)"  
/protein\_id="AA082422.1"  
/db\_xref="GI:1086831"  
/translation="MADYAVARRSAPTGRVLYKAIETGFRGLRTQSESTLKLKE  
GVFNKEDCFAGKRVLYLYKAHKKTLTGHTVTFRAIWKITRPHGNAGVARAF  
HHNIPSPSLGRKRVILYPSNI"  
complement(5858..9377)  
/gene="F10E7.9"  
/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=F10E7.9;class=Sequence>"  
complement(join(5858..5968,6012..6221,6265..6633,6769..7271,7551..7752,7801..8020,9268..9377))  
/gene="F10E7.9"  
/note="coded for by the following C. elegans cDNAs: cml391"

## gene

## CDS

/codon\_start=1  
/product="Hypothetical protein F10E7.9"  
/protein\_id="AA082423.1"  
/db\_xref="GI:1086832"  
/translation="MSEGTGSSSHSVVPEPDSIFSGEETADSSQLPMWORTILTRDK  
VLFGTGVGFATVWNIIFGLIFRLMIYGTAVANVSLILGCTSLALITVYSAIG  
IVERCOIKSGGIYFLVAVHVGIOIGAGIITVAVAGQAVVAVGSEVANHFESE  
SKMIKGIATITLAVLTAVNTAGTYTAVRQIYVLLITLALAVDFITGALPSSPEEG  
VFRSSSEKIRYNASHYEAVNCSITIGIPROIPESFTVFGVFANLGLVAGVNMKG  
DLKDPKHSIRPLGELSAVGSSTICFIFIMLGGDMFLCDPMANLSEKVALGIVEL  
VGLYVCSLSVGPNDNPVLAVGLMCAVSFLVGLDNLALISTMPALITVAYVY  
SYVLSAYSDATQVHADDEVETVGTGKNKNDNELEPFRDYNITSTETETGYS  
QEATVWSEMSRNVSPICFLVNAIILLIHPVAVAFVAILIAYWYIGVCPGCGG  
ISTEISLCHMRKTVYSSLDLSRGATTTLPETTFISGVSSSQQNEANPDYDROKRYHHS  
ENVSQHLD"  
11359..12212  
/gene="F10E7.5"

## gene







TITLE	Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium
JOURNAL	Science 282 (5396), 2012-2018 (1998)
MEDLINE	99069613
PUBMED	9851916
REMARK	The <i>C. elegans</i> Sequencing Consortium.
REFERENCE	2 (bases 1 to 34368)
AUTHORS	Basham V.M.
TITLE	Direct Submission
JOURNAL	Submitted (06-NOV-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jlesaner.ac.uk or rwememate.wustl.edu
COMMENT	On Nov 29, 1997 this sequence version replaced gi:1695114. Coding sequences below are predicted from computer analysis, using predictions from GeneFinder (P. Green, U. Washington), and other available information. Current sequence finishing criteria for the <i>C. elegans</i> genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is not the entire insert of clone W06D12. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone W06D12 is at 553 in sequence A1023857. The true right end of clone W06D12 is at 2303 in sequence Z93782. The start of this sequence (1..104) overlaps with the end of sequence A1023857. The end of this sequence (34265..34368) overlaps with the start of sequence Z93782. For a graphical representation of this sequence and its analysis see: <a href="http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=W06D12">http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=W06D12</a> IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. Location/Qualifiers
FEATURES	1..34368
source	/organism="Caenorhabditis elegans"
gene	/db_xref="taxon:6239"
gene	/chromosome="V"
gene	/clone="W06D12"
gene	join(1055..1180,1224..1326,1402..1631,1673..1816)
CDS	/gene="W06D12.6"
CDS	join(1055..1180,1224..1326,1402..1631,1673..1816)
CDS	/gene="W06D12.6"
CDS	/codon_start=1
gene	/protein_id="CAB04927.1"
gene	/db_xref="GI:3880551"
gene	/db_xref="SPRMBL:045895"
gene	/translation="MEOLETCDILKFTPNSPENTEMRYVETIPGVSNLADYLSIGPQFTMAGYVCAKGMHNGFALFVPEPKSSITRHCNCDPVSYVTFYQHSPTNQFTRAIRDVYINRTIEGTRKNVSECFYTCAGDDLVFVSWENGPTIVRFPGQITACRPRPVSHSDVMKPFYENWFMFDMYMOIPOITCFE"
gene	join(2263..2429,3400..3736,4267..4764)
gene	/gene="W06D12.7"
CDS	join(2263..2429,3400..3736,4267..4764)
CDS	/gene="W06D12.7"
CDS	/codon_start=1
gene	/protein_id="CAB63322.1"
gene	/db_xref="GI:6580250"
gene	/db_xref="SPRMBL:090346"
gene	/translation="MTSSIMNENPFYIMTROLIIFGLIFELFYPLIFIVYLCRTPKQCNLKRKSFHFHFAATCOMLCSFLIIDFTYVPSPKITIIDGILIDIFDAKIVYTSMTILSALTSSELCFCSVFNIFEMLNKNSWLTSEEMLMTVLMFPFMSLIGFTIRQDPNVAKLTKQISSIYPCIDINPNVILPILPAEGKSLVEFLMLFTMAASLISADLISHLRMTNOSDKTRKMKHOKFNRSRLFOVLIDISTSPFYIISNFATLFWOMPELTYPDFMSKSGSTACIFVIFMYVDPYRKMLQEIYARSKSTQTSVRHVTIIVADRGNRP"
gene	complement(join(5022..5258,5402..5535,5584..5957,6431..6584,6629..6770,6823..6945,7344..7618,7718..7908,8766..8883,9010..9145,9196..9207))
CDS	/gene="W06D12.5"
CDS	complement(join(5022..5258,5402..5535,5584..5957,6431..6584,6629..6770,6823..6945,7344..7618,7718..7908,8766..8883,9010..9145,9196..9207))
CDS	/gene="W06D12.5"
CDS	/note="cDNA EST yk224a2.3 comes from this gene"
CDS	cDNA EST yk224a2.5 comes from this gene"
CDS	/codon_start=1
gene	/protein_id="CAB04926.1"
gene	/db_xref="GI:3880550"
gene	/db_xref="SPRMBL:045894"
gene	/translation="MTSEQNNYTPRGNNPFRSAVLBEIHEEDSDTSPEEDTLENKPSHENDFKRNALRSEETPEISPEIORIEDPLASDSEENKFPASISVSNADLERHRDLSIPENDRSQSYTRHSIDKGLIKQNNSEVDVEYERNNYVGRPRPFYNKLPGRLOHRLSVHTDASRFSFGAMGVPTLRPFKSNWFAFHRKQIGFRHSVVILVILYTLTGAVMEFTVSSRHEKAKTLDHVNLIHLDRLAENITESSYNNITTTTDEEMKYIRATYELMKLBQYKSTYYKLEADNNKWTRESAFPSSMNYTTTGGSIAPESTLGOVLVCVGFIEVPTLVVLRDGFELVHLTKLVHGIQKRELNKHHVDEDEIISLPKACLLLASYLACTIEIFYDELSPEPTQMDMLCEYFSISLSTIGLGDIMPNNATVGNRNNTYENKIFAPISIIIFEGMAVTVVRNMEYIAVENGIFGATLVENKLDALVTSSASVREDEPSPTRKQVRLASVDAGSEODETPHEINLNTLRSIATFMKSMADLYGGGFGVQLRRDLIHSDSQNMQTMVSMQRLHRRASQSQPQA TGCNNNV"
gene	join(13888..14185,14708..14777,14968..15102,15147..15423,15938..16234)
CDS	/gene="W06D12.4"
CDS	join(13888..14185,14708..14777,14968..15102,15147..15423,15938..16234)
CDS	/gene="W06D12.4"
CDS	/note="contains similarity to Pfam domain: PF01461 (7TM chemoreceptor), Score=36.9, E-value=5.1e-13, N=1"
CDS	/codon_start=1
gene	/protein_id="CAB04925.1"
gene	/db_xref="GI:3880549"
gene	/db_xref="SPRMBL:045893"
gene	/translation="MTVSFEVSDNSYLIVITCVSCNILLILVYTKSKRHGTXYLMFPLANGILHSVFQDITPILMLDGNKHSILVFLNKRFSKMAVLASISYGFELAVIFAVHFIETRWITIKNDKIRFENPFIQWIGSYGVEYSYLHVEFSERPDKTEFIRKQMSDFGHTMDEITVMARFVSEEGALIDIRQGVFYAAKRTMLSPILLIYCGAIIKTRTKRLDKLSRKRLELTENLALIVQVIVPLMNYCFLMMNEPFLIGETARLTGMVGLIPGIDILAIIFIVANRRALKRFRVELVQPSSTVRRDISTSRALAQVAGNOYAPAVVIN"
gene	join(19946..20159,20215..20551,21208..21658)
gene	/gene="W06D12.3"
CDS	join(19946..20159,20215..20551,21208..21658)
CDS	/gene="W06D12.3"
CDS	/note="contains similarity to Pfam domain: PF01069 (paty acid desaturase), Score=465.8, E-value=1.1e-136, N=1"
CDS	cDNA EST yk148b12.5 comes from this gene
CDS	cDNA EST yk412h2.3 comes from this gene
CDS	cDNA EST yk412h2.5 comes from this gene
CDS	cDNA EST yk548c4.3 comes from this gene
CDS	cDNA EST yk197g12.5 comes from this gene
CDS	cDNA EST yk548c4.5 comes from this gene
CDS	cDNA EST yk661f6.5 comes from this gene"
CDS	/codon_start=1
gene	/protein_id="CAB04924.1"
gene	/db_xref="GI:3880548"
gene	/db_xref="SPRMBL:045892"
gene	/translation="MTQIKVDYIISKQFLADINETROMDESKKOVIKEIYKQVVALHIGALVGLYQLVFQAKWATVGVFLHTLIGSMGVGGAHRLMAHRAVATLSMRFVLMILINSIAFONDIDVARHCHHKWTQDADPSTIRNGHFAFMGLVYKRDOLKIOGKLIADSDLEDYEDPVLFRKRNKLLPVIGICFALPPIPVLVNGESFIAPYAALEFYCFILHATWCINSYHWGQOPYDHOSSNDWTMTSIAAGEGSHNHFPPDYRTSEHAPELNMTPFVLIDFGASIGMVYDRKTPREVIOCKKSGCEPERRMKHKG"
gene	complement(join(21764..21883,22445..22683,22761..23145,24058..24140,24197..24335,24883..25051,25103..25312))
CDS	/gene="W06D12.2"
CDS	complement(join(21764..21883,22445..22683,22761..23145,24058..24140,24197..24335,24883..25051,25103..25312))
CDS	/gene="W06D12.2"

```

/gene="W06D12.2"
/note="cDNA EST yk10198.3 comes from this gene
cDNA EST yk10198.5 comes from this gene"
/codon_start=1
/protein_id="CAB04923.1"
/db_xref="GI:3880547"
/db_xref="SPTREMBL:O45891"
/translation="MFNVLAVAVYERHHLHYKIGVLYSFLGAGLYLCEAENEK
SLKHEDNMRVRLSLIAKOVFORLQNMVSGNSSESESTKRLRRLVSEDAAMGSI
DSKMKTRMDIWMGLYAGITTYTIGYDLAIGTIGWICITMLYAMIGIPVILNDW
GNMIFEVDFHFWONIGROMIOSLROKLRKRVOSLESGSIDKPLMETSTPSPND
NGTRPIPLLVLLVLEFWMIOCVAFAYFENNTLPSYVEFFISMTTIGDPTSPHT
VAVGIVFIIIGLSVSMCIINVLOMLEFTFNQIQRINDPKNTLSVAEERKRSI
GVSEPSIDPSKKKMTMDGNVNAEKISENNEKMKLMRPMNHQKMKLNEKFERAK
MNSGTOTTSQIVASVQTAERYEQTWDEDEEOPKRSINTRLTYNTGE"
complement(join(28517..28645,28743..28885,28938..29044,
29096..29211,30024..30152,30200..30467,30519..30661))
/gene="W06D12.1"
complement(join(28517..28645,28743..28885,28938..29044,
29096..29211,30024..30152,30200..30467,30519..30661))
/gene="W06D12.1"
/codon_start=1
/protein_id="CAB04922.1"
/db_xref="GI:3880546"
/db_xref="SPTREMBL:O45890"
/translation="MQPASELPDSIQQLNDSDAPVCKYEVLMEQAQPLSHATVGL
VHKMSCDNNKEMCYMTVHSCVVDGQFGOKLYDEHCTTLDALILKELEYNKLE
AGOMSVFEPKADKPIYFSCMIKVEKESAEPCVYFCEKRLPESKQNDLVKLE
KVSDDIPGEPQONTNKLQSKQESTDAMTSQVSAFLELDELDEVDKMSRPSV
TMNRLIRDVAGNTKTLADVPAASVAVODPESENGKRNRFHVPNTASNDVC
FPRNLIISLLILAAIGIYIISIMYLSLOVGRKPNKRSITFDEQGYFNFYTGIM
ICSEFVILIR"
complement(29577..29648)
/gene="W06D12.t2"
complement(29577..29648)
/gene="W06D12.t2"
/note="TGA 777 2-tRNA
predicted using tRNAscan-se-1.11
preliminary prediction
similar to tRNA-777"
29642..29714
/gene="W06D12.t1"
29642..29714

gene
trna

Query Match 4.5%; Score 52.8; DB 3; Length 34368;
Best Local Similarity 53.6%; Pred. No. 0.075;
Matches 133; Conservative 0; Mismatches 112; Indels 3; Gaps 1;

```

ACCESSION 278542  
 VERSION 278542.1 GI:1507784  
 KEYWORDS HTG; ADP-ribosylation factor; Cuticlin; Human mRNA KIAA0066 protein  
 SOURCE like; Transcriptional activator protein.  
 ORGANISM Caenorhabditis elegans.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE  
 1 none.  
 2 genome sequence of the nematode *C. elegans*: a platform for  
 investigating biology. The *C. elegans* Sequencing Consortium  
 Science 282 (3596), 2012-2018 (1998)  
 99069613  
 MEDLINE The *C. elegans* Sequencing Consortium.  
 2 (bases 1 to 40600)  
 REMARK Burton J.  
 AUTHORS

TITLE Direct Submission  
 JOURNAL Submitted (23-NOV-1996) Nematode Sequencing Project, Sanger  
 Institute, Hinxton, Cambridge CB10 1SA, England and Department of  
 Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:  
 jess@sanger.ac.uk or rwnematode@wustl.edu  
 Coding sequences below are predicted from computer analysis, using  
 predictions from GeneFinder (P. Green, U. Washington), and other  
 available information.  
 Current sequence finishing criteria for the *C. elegans* genome  
 sequencing consortium are that all bases are either sequenced  
 unambiguously on both strands, or on a single strand with both a  
 dye primer and the terminator reaction, from distinct subclones.  
 Exceptions are indicated by an explicit note.  
 For a graphical representation of this sequence and its analysis?  
 see: <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=F20D1>  
 IMPORTANT: This sequence is NOT necessarily the entire insert of  
 the specified clone. It may be shorter because we only sequence  
 overlapping sections once, or longer because we arrange for a small  
 overlap between neighbouring submissions.  
 IMPORTANT: This sequence is not the entire insert of clone F20D1.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we arrange for a small overlap between  
 neighbouring submissions.  
 The true left end of clone F20D1 is at 1 in this sequence. The true  
 right end of clone F20D1 is at 4761 in  
 sequence AL031620.  
 The true left end of clone C18B12 is at 40500 in this sequence. The  
 start of this sequence (1..104) overlaps with the end of sequence  
 Z83227.  
 The end of this sequence (40500..40600) overlaps with the start of  
 sequence AL031620.

## FEATURES

source  
 1..40600  
 Location/Qualifiers  
 /organism="Caenorhabditis elegans"  
 /db\_xref="taxon:6239"  
 /chromosome="X"  
 /clone="F20D1"  
 complement(join(3587..3770,4128..4253,4308..4427,  
 4490..4560))  
 /gene="F20D1.1"  
 complement(join(3587..3770,4128..4253,4308..4427,  
 4490..4560))  
 /gene="F20D1.1"  
 /note="predicted using GeneFinder  
 cDNA EST EMBL:Z14514 comes from this gene  
 cDNA EST yk61b9.3 comes from this gene  
 cDNA EST yk60a5.5 comes from this gene  
 cDNA EST yk61b9.5 comes from this gene  
 cDNA EST yk327b8.3 comes from this gene  
 cDNA EST yk327b8.5 comes from this gene  
 cDNA EST yk314h2.5 comes from this gene  
 cDNA EST yk328d5.5 comes from this gene  
 cDNA EST yk335d5.5 comes from this gene  
 cDNA EST yk413f2.5 comes from this gene  
 cDNA EST yk416d2.5 comes from this gene





DEFINITION Caenorhabditis elegans cosmid K06A1, complete sequence.  
 ACCESSION U23449  
 VERSION U23449.2 GI:14574206  
 KEYWORDS HTG.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Caenorhabditis elegans.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 1 (bases 1 to 26449)  
 Waterston,R.  
 TITLE Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium  
 JOURNAL Science 282 (5396), 2012-2018 (1998)  
 MEDLINE 99069613  
 PUBMED 9851916  
 REFERENCE 2 (bases 1 to 26449)  
 AUTHORS Fulton,L.  
 TITLE The sequence of C. elegans cosmid K06A1  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 26449)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-JUL-1995)  
 REFERENCE 4 (bases 1 to 26449)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-JUN-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 5 (bases 1 to 26449)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JUN-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 6 (bases 1 to 26449)  
 Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 On Jun 28, 2001 this sequence version replaced g1.733545.  
 Submitted by:  
 Genome Sequencing Center  
 Department of Genetics, Washington University  
 St. Louis, MO 63110, USA, and  
 Sanger Centre, Hinxton Hall  
 Cambridge CB10 1RO, England  
 email: rwenemate@wustl.edu and jessesanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

For a graphical representation of this cosmid sequence and its analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=K06A1,class=Sequence>

NEIGHBORING COSMID INFORMATION

The 5' cosmid is K03H9, 900 bp overlap; the 3' cosmid is T28D9, 2500 bp overlap. Actual start of this cosmid is at base position 897 of K06A1; actual end is at 16582 of T28D9.

NOTES:

## FEATURES

## source

Coding sequences below are the result of integration and manual review of the following data : computer analysis using the program GeneFinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuiji Kohara ([http://www.ddbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and The C. elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>), similarity to other proteins from Blastx analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's MAM alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

## 1. 26449

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

/db\_xref="taxon:6239"

/chromosome="II"

/clone="K06A1"

98..3998

/gene="K06A1.3"

/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=K06A1.3,class=Sequence>"

join(98..182,258..358,462..536,656..798,896..1277,1325..1780,2763..3198,3244..3494,3687..3773,3819..3998)

/gene="K06A1.3"

/note="similar to C. elegans cuticulin precursor; coded for by the following C. elegans cDNAs: YK334a2.3, cm20c2, YK60c1.5, YK334a2.5"

/codon\_start=1

/product="Hypothetical protein K06A1.3"

/protein\_id="AAC24298.2"

/db\_xref="GI:14574207"

/translation="MTSTDFLLMLALFNSAQPIVEEDNHSIDPEVSGHSGFSLKYNTEKSPSHVFKGHRKDGCSFNTAATPEFSKCDVMRORANPKMAYTAVVQ LHPITPTVDAVYKLRCPKAEKAVGAEVSDPTPVQLDESPQVCSYTIKESPNGLAKFAQDVLVHVEGCESEAVOMVGVGGEYKKGKVGEGCESEDIYIMP NLIVENRTKAFVNSNAPFPDQNNVRSCTISGTCGCHPKDASASTDDAS EPIGEELARDELITPMPTSSSEITPTSTTTTSTTTATETPTTTERPTTVQ VLDPTTITLPTDPTDSTIRFLSKPSKSNNTWEDSGEITSEBSHAHAQHM TSSGPTTKREIKRRDQATVDDEHROISYLAASVLADVILNARCFVNSLKLFL LKGVFYVTFSSRHAGRPRTGOMOPFHYVGLYLRNSKRSISSGNCINNOGNNQ CTSSGCFHFATSEGDIERFDCSOPKSLSTLCAVLNKTDESSVCLGARPTHC CTSSSPKICAPKRRKMAELQMLTMSQVWMLVGHVGLSCVVLIVFYWKSVKYL PLYVIPEESRMTYITLAKIKFDAGIGSLYSPFLIFPIPEFSFIYKDKKINKTC ALFLAIFKLTIVPIFYECFLSTVCLMHKDHDPGICRLNGAIIWQFMVFR"

4064..4803

/gene="K06A1.2"

/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=K06A1.2,class=Sequence>"

join(4064..4237,4286..4362,4415..4490,4535..4599,4644..4803)

/gene="K06A1.2"

/codon\_start=1

/product="Hypothetical protein K06A1.2"

/protein\_id="AAC24297.1"

/db\_xref="GI:733347"

/translation="MLLVITQVYLLSKURPNIDIPRISEETPSQVAPRTIISGEGTR IESEFENETVYIHPRPFPVRLGISILWKNRNNMDLALRLALSTGFLYKLPSEFLI PGKEISIEVQCTRAESPILNHNILTEMFLDSTPTPGDLORLMSKPYLVPOSHMXY FVLVYVDDSTVSPSHVSSVE"

complement(5030..8457)

/gene="thr-22"

/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=K06A1.4,class=Sequence>"

complement(join(5030..5039,5091..5185,5232..5383,5430..5545,5594..5725,5773..5969,6018..6415,6852..6983,7072..7234,8113..8457))

/gene="thr-22"

## gene

## CDS

```

/note="similar to C4-type zinc finger domain (exon 2);
coded for by the following C. elegans cDNAs: YK325g8.5,
YK324h4.5, YK324h4.3, YK325g8.3"
/codon_start=1
/product="C. elegans NHR-22 protein (corresponding
sequence K06A1.4)"
/protein_id="AAC24301.1"
/db_xref="GI:3253111"
/translation="MSMDLIEYFYPPTDLFRNFEDVPPSPSLDVGDLSGSSPYD
LGFQPSFDIDFCIWKGGSSSPPESSNHLPEITVKAISTSPGMPSPILDSKCHVSP
TANTLHFHGGSCACAAFFRRYSMSMTICIGTGDTPNCFRHYELRMICRCHRIK
CIDAGNRELVQARKKEDEYRRKSGLVYKEDDVESVDEYVNTYNNVODSSPK
MESGINENTIPESYLSPPSSQPLDMVTYPPPLSTPESILTPATGDMHCORVIS
NSPPDIFHASTGSHAGFIOHVQFVPEVANKTEELVDHYHREASLNBRRKIMYDT
OIRDFVDTCECEYENHOLKRFYKSCFQVKNDFYMIIDYVQFEPFQALHNDKAV
VTRMACAVDSMLASAYSTKVGTEKRLIFNGDIYNNMPTISGDEPGAGTEFQPO
EHEKRYTLPMLKQYFDLAIIPARLEVEFEVLLKALIMQISNRYLLEERAIICA
ROHDIYVQALHKKVVEBERGDEDPAIRVGOLLSMSYTEQVQAMTNSYLVMTFEDVVC
DSIMYDLSEFRD"
complement(9832..11847)
/gene="K06A1.5"
/note="for a graphical representation of this gene see:
http://www.wormbase.org/db/seq/sequence?name=K06A1.5;class
=Sequence"
complement(join(9832..9922,10394..10716,10769..10943,
10993..11636,11686..11847))
/gene="K06A1.5"
/note="Contains similarity to Pfam domain: PF00400
(G-beta), Score=70.8, E-value=9.6e-18, N=5"
/codon_start=1
/product="Hypothetical protein K06A1.5"
/protein_id="AAC24299.1"
/db_xref="GI:733550"
/translation="MANVYRMKSKNDQDIDANRKLADSESRYSVQDRKLRIFED
AIVKMTVLENELAEIKENSVINTERVSLVATCNLTKEKTKOLDNERQLNRIEL
OEKSAEFMAEALIOEERAOMRIEOIAKATADNLNGERASSAFGSPETLDEPMAT
DVPSEYFKELSTHDEGHVEMWSDPFAATAGSDKVOIMRYSPKTKASKYTSIG
CICPVNRIDYDSORHYCLASNDKRCRKNIDSORLSTFSFGITDVSSRLQSHNV
ISSADRTIKNMDISSIRCLKSTLVGTFVFDIYAKGVSSVSISSHDFKVFVWAR
SSDQATYVELGQVSSLDISMDGLQVLAASRDYTLIDVRNNGIHLVSAEQYKTD
DSTRALFSSSTGEVLAGSSSVFIMNTYTKTLKLVKVAKASDSAOIMSLANMPSGRG
LIACDRQKTCCTLMR"
13643..15109
/gene="K06A1.1"
/note="for a graphical representation of this gene see:
http://www.wormbase.org/db/seq/sequence?name=K06A1.1;class
=Sequence"
join(13643..14066,14112..14328,14377..14495,14671..14939,
14993..15109)
/gene="K06A1.1"
/note="Similar to transcription factor AP-2
(SP:7AP2_HUMAN, P05549)"
/codon_start=1

```

Query Match 4.2%; Score 49; DB 3; Length 26449;  
 Best Local Similarity 59.9%; Pred. No. 0.59; Mismatches 55; Indels 0; Gaps 0;  
 Matches 82; Conservative 0;

```

271 AATCCAGGTGATATTTTGTACAAACAACGTGTGTCATTTGTTTCATTCATTAATTTGTT 330
      |||||      |||||      |||||      |||||      |||||      |||||
Db    659 AATCCGAAAGAAATGCGATATACAGCTAGTACGTTCATTCACCTCCTTTTACA 718

QY    331 ACCAAAGTTATGTCGATATCGAGTACATGCTTTTACATGGAAGCTGATAAACAGTT 390
      |||||      |||||      |||||      |||||      |||||      |||||
Db    719 ACTAAAGTGCAGAGCGCTATAACTGCTGTTTCTACAAAGAACCGAAAGCAGAGTC 778

QY    391 AGTGCACAGATTGAGT 407
      |||||      |||||      |||||
Db    779 GGTGCTGAGGCTTTCTGT 795

```

RESULT 33  
 U21318

LOCUS U21318 31731 bp DNA linear INV 23-MAY-2002  
 DEFINITION Caenorhabditis elegans cosmid K03H9, complete sequence.  
 ACCSSION U21318  
 VERSION U21318.1 GI:687828  
 KEYWORDS HTG.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Caenorhabditis elegans.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 1 (bases 1 to 31731)  
 Waterston,R.  
 Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium  
 Science 282 (5396), 2012-2018 (1996)  
 99069613  
 9851916  
 2 (bases 1 to 31731)  
 Anderson,K.  
 The sequence of C. elegans cosmid K03H9  
 Unpublished (2001)  
 3 (bases 1 to 31731)  
 Waterston,R.  
 Direct Submission  
 Submitted (13-JUL-1995)  
 4 (bases 1 to 31731)  
 Waterston,R.  
 Direct Submission  
 Submitted (23-MAY-2002) Department of Genetics, Washington  
 University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
 Louis, MO 63110, USA  
 Submitted by:  
 Genome Sequencing Center  
 Department of Genetics, Washington University  
 St. Louis, MO 63110, USA, and  
 Sanger Centre, Hinxton Hall  
 Cambridge CB10 1RQ, England  
 email: rtw@emslab.wustl.edu and jess@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we provide a small overlap between  
 neighboring submissions.

This sequence was finished as follows unless otherwise noted: all  
 regions were double stranded, sequenced with an alternate chemistry  
 or covered by high quality data (i.e., phred quality >= 30); an  
 attempt was made to resolve all sequencing problems, such as  
 compressions and repeats; all regions were covered by sequence from  
 more than one m13 subclone.

for a graphical representation of this cosmid sequence and its  
 analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=K03H9;class=Sequence>

#### NEIGHBORING COSMID INFORMATION

The 5' cosmid is T25D10, 200 bp overlap; the 3' cosmid is K06A1,  
 900 bp overlap. Actual start of this cosmid is at base position 1  
 of K03H9; actual end is at 8345 of K06A1.

#### NOTES:

Coding sequences below are the result of integration and manual  
 review of the following data: computer analysis using the program  
 GeneFinder (P. Green and L. Hillier, personal communication), the  
 large scale EST projects of Yui Kohara  
 ([http://www.ddbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and The C.  
 elegans ORFome cloning project (<http://wormdb.dfci.harvard.edu/>),  
 similarity to other proteins from Blastx analyses  
 (<http://blast.wustl.edu/>), sequence conservation with C. briggsae  
 using Jim Kent's WABA alignment program (Genome Research  
 10:1115-1125, 2000), individual C. elegans GenBank submissions,  
 and personal communications with C. elegans researchers. tRNAs



similar to SWISS-PROT:P34029 (DP3B\_SPIC1) Blastp2 p-3e-14  
C-43%"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="DNA POLYMERASE III, BETA CHAIN"  
/protein\_id="CAC13175.1"  
/db\_xref="GI:14089415"  
/translation="MKIFLEKSRFLRPLENTTKTIISNSNSNYELRCAYFVYRQDKIIL  
ISSNEDISTINNEYLLDGIKNTKGESDFVLNSLKNIIKSCDSIIILEKEMTKLIF  
SDLSQYEINLILDEHREPDIKFGNNSQLQKTEQFQALKNVFPADLSNAOELLIS  
VNLSTKNMILTVAITNARIAQMOKITEDSQEENLITNSKVKEILSSMNTLISP  
GPELKISGNLEIKTKVIEIPYMANVENFEDQENFVHIDKKEILSLIDKSVLDE  
KNGNKILIEVNSPKSKKIKLSSYWDIGFSEVSDNEFEVSEELKLFITNAVYAKES  
INVFDMISFIETENKDRWYISSETNLNKKLIAALRGH"  
2864. .3085  
/gene="MYPU\_0030"  
2864. .3085  
/gene="MYPU\_0030"  
/note="identified by sequence similarity; putative:  
similar to SWISS-PROT:P05650 (YAA\_BAC50) Blastp2 p-6e-06  
C-45%"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="conserved hypothetical protein"  
/protein\_id="CAC13176.1"  
/db\_xref="GI:14089416"  
/translation="MTILKGEFLITGOLIKIKLIDSGQIKQFMOSHQVINGKLI  
MTRNTYRVDITWIDQLYTIKQNNEE"  
3088. .3771  
/gene="MYPU\_0040"  
3088. .3771  
/gene="MYPU\_0040"  
/note="identified by sequence similarity; possible:  
similar to SWISS-PROT:P43055 (YL1\_MYCHO) Blastp2 p-1e-28  
C-53%"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="conserved hypothetical protein"  
/protein\_id="CAC13177.1"  
/db\_xref="GI:14089417"  
/translation="MKYIKSKIHKKVPRKLSHYKLFYGFPHNTYIGFYNGIVQIRL  
KINDIYNHEDEYKYLKNNEDYLVNKFVYTKKPDGHEFKANKESLKNLDIJEKES  
QNDJLAKMMLFYTKRDPKYLSLVERKDFNDVHGDINFKMILANDKNEKVLIDFE  
WIRKGSYEDIIICYKNFNINKNELIKRDLNEKEFDYLYMSVKFEIEAEKERYSKM  
SDSFWEYERK"  
3778. .4422  
/gene="MYPU\_0050"  
3778. .4422  
/gene="MYPU\_0050"  
/note="no similarity found"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="unknown; predicted coding region"  
/protein\_id="CAC13178.1"  
/db\_xref="GI:14089418"  
/translation="MERNVSEFVQICEKNLFLYSLAFETLSMNGEKIDOKKHIIQ  
VFMSIDSYRLKKKFNVNVIDNSINFLKRPVLDKSNMINEVSEFVILIPST  
LKKIKSPKSLKMLFIOINKNKEIKNSSLFKKIPNFPINIKAIATSYSSKHGCGF  
VLSNVNENTKKNFLNITFKTELQFLNLTATRYIEELIESHFKNRYKDGIKITF"  
4603. .6759  
/gene="MYPU\_0060"  
4603. .6759  
/gene="MYPU\_0060"  
/note="no similarity found"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="unknown; predicted coding region"  
/protein\_id="CAC13179.1"

/db\_xref="GI:14089419"  
/translation="MKKSPKRNPFKSSKYVDKRTKRNITIGSIFTLIVGVVPIE  
AVNSTAQRPALPEDAPLNFINGRNSDQVTSYDVSILNGVLRKEVKKQKQEVFTKSN  
AYSEFLYDGRKASENLANFVKVESGSPSEFTNLKRIIDETIRASERELDAKRN  
IQLOOGPANAEAWLKQIRENDKGGKGNEDALEFVQKRIKNDVARRDPQILISL  
NVDLKKVASRDIDHDNTKVFLEKQOPPEFADFTNKQVNFPSNDSNTKSAV  
YTRRSFIFNEHRNENLAKRYLERGKTIYSLTILPKSNDARKEEYVSKDFTL  
ISYSAFNPDLQDATTNNNGVDLILKANKNYGVEFLPSNISAEPERQADIDHITN  
VASYNSKNGRGVYRGSLDVSTKTSIDLSTLTIGITFVVAELDATNINPLDTFIE  
KVENFEGLLPDAPANVSADAKTNIRELSOKIOWSKKEEPIKRIKEAVETTPKRN  
GREILQKRMAYYYKAKIDTDNENYVDDGLNIVKKEHSTYDEFWNYLKNLYSL  
HSSSTORVGEDPDKLKYSDNPFPAQGFSAEFKPYREKYSNLTVMNFSERN  
ARAFELNINNHQSNOSKVEKIRKFIESKKNYVELDENIEFLKATPNDINIK  
OSPREKVDSKSFEEIILPMFDFVGSRLKQDN"  
6759. .8189  
/gene="MYPU\_0070"  
6759. .8189  
/gene="MYPU\_0070"  
/note="identified by sequence similarity; putative:  
similar to TREMBL:Q49557 (Q49557) Blastp2 p-3e-15 C-298  
piscan: pos. 18-28 PS00013 I PROKAR\_LIPOPROTEIN  
prokaryotic membrane lipoprotein lipid attachment site"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="p60-like (Mycoplasma hominis) LIPOPROTEIN"  
/protein\_id="CAC13180.1"  
/db\_xref="GI:14089420"  
/translation="MINKIKIKILKTFPTSTPIITSSILISGQVIDSTERKINDET  
KITSKESVNLVYEVYISQVSNDSIDFSTKNDSLODKNSQFKAFLAKFYVN  
ERLKSSTWPNYDORANMEKANTSFVNKTOSIJTMNGHPLDENFVLNKTSPYR  
HEITKMLMKNVYLLNTSEDOVKKSFPEYKRIKILDKNNDKRTYESIDSKNPFEL  
KYLKDKRMIMKMYKSETDSSKVLADIVYINIOFENENVOFOSVESLITBEDISL  
RLNDEYKRAKETQNSRLNPELIRAFETVETTEAGGLANIIDLOKQKEVINGFLVD  
KNKISKQALLINQASQSGDGLTRFQIOTNFTKAGKSEIMSYLTKLPPEEDTAT  
AEVYKDKDSDPTIKRGCFTEFKFNPMQDAESANELTKLAVWLYPRDNLFTQAKNH  
FSSIGYKIESNNPIIRDLKNEGILK"  
8190. .8540  
/gene="MYPU\_0080"  
8190. .8540  
/gene="MYPU\_0080"  
/standard\_name="hlt"  
/note="identified by sequence similarity; possible:  
similar to SWISS-PROT:P32083 (YHIT\_MYCHR) Blastp2 p-7e-22  
C-368"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="HIT-like PROPEIN (CELL CYCLE REGULATION)"  
/protein\_id="CAC13181.1"  
/db\_xref="GI:14089421"  
/translation="MSDNHSIFKKIIDRKEPAQIYEDDOAIAIMDKPPYNPCHFLV  
IPKHSITNLKDIEOSINHLKIAIKLAKKIEKFEDEFKLIITNNEKAGOVVYHTH  
IHIFPRKKEGHEH"  
8643. .9296  
/gene="MYPU\_0090"  
8643. .9296  
/gene="MYPU\_0090"  
/standard\_name="rplK"  
/note="identified by sequence similarity; possible:  
similar to SWISS-PROT:P75550 (RUL1\_MYCPN) Blastp2 p-6e-36  
C-448"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=4  
/product="50S RIBOSOMAL PROTEIN L11"  
/protein\_id="CAC13182.1"  
/db\_xref="GI:14089422"  
/translation="MEVDQWLTLPHILERIVILMAKKEVYKIAQLQFIAGKAPGPS

Query Match 4.2%; Score 48.4; DB 1; Length 327650;  
Best Local Similarity 46.7%; Pred. No. 0.77;  
Matches 154; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY	108	TTTCATTAACAAATCAATTTTAAATACACGTAATGCATTCGAAAGGACATGTTTAACTGCAAAAG	167
Db	52535	TACAAAAAATAAATGATGATTAACATACATGGATACACCTCAATATCTCTGCTTTTCT	52594
QY	168	TCTTTTATGATCAAGAAGTTGCCGTAATCATGAAGTGCACGTCAAGTTGCCGGAATTC	227
Db	52595	TGGCTAGATTTTACCTGATTTAAGTAATATGAAATTTATGCAAGCAAGGAGGAGGTTTC	52654
QY	228	ACTTCATTTTATTCATTCGCAATGTTTCCGGTACACGATCTCGAATCCACGAGATATTT	287
Db	52655	AATAGTATTTAGTATCATATATATATGACATTTTCTATTAATTCATACAAATTTGTTTTT	52714
QY	288	TGTACACACACTGTTTGATCTTTCGTTTCATTCATTAATTTGTTACCAAGTTGATCGTGC	347
Db	52715	TGTTCCAAATTTTGTGTTTATTTGCTTTTGGCAAAAAAATAACTTTTCAATTAATCTAC	52774
QY	348	AATATGAGTACAAATGCTTTTACATGGAACCTATAAAACAGTATGCGACATTTGAGGT	407
Db	52775	AATGTTTTTTTGAAGGATTTTCAATTAATGACATAAGTATTTTGTATTAATTTACATCTTTC	52834
QY	408	ATCTGAAATACACACTGCTTTTCAAACTCA	437
Db	52835	ATCTAATAGTATTCCTTCAATGACAAAGTAA	52864
RESULT 35			
LOCUS	AC116984	53150 bp	DNA
DEFINITION	Dictyostelium discoideum chromosome 2 map 3000630-3053778 strain AX4,*** SEQUENCING IN PROGRESS ***, in ordered pieces.		
ACCESSION	AC116984		
VERSION	AC116984.1	GI:20042947	
KEYWORDS	HTG; HTGS; PHASE2.		
SOURCE	Dictyostelium discoideum		
ORGANISM	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.		
REFERENCE	1 (bases 1 to 53150)		
AUTHORS	Gloeckner,G., Elchinger,L., Szafranski,K., Pachebat,J., Dear,P., Lemann,R., Baumgart,C., Parra,G., April,J.F., Gulgo,R., Kumpf,K., Tüngel,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.		
TITLE	Sequence and Analysis of Chromosome 2 of Dictyostelium		
JOURNAL	Unpublished		
REMARK	The Dictyostelium Genome Sequencing Consortium		
REFERENCE	2 (bases 1 to 53150)		
AUTHORS	Baumgart,C.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (05-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany		
AUTHORS	3 (bases 1 to 53150)		
JOURNAL	Biotechnology, Beutenbergstr. 11, Jena 07745, Germany		
REFERENCE	Baumgart,C.		
AUTHORS	Submitted (18-MAY-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany		
JOURNAL	CDS predictions from GenBank may contain errors. Further information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/)		
COMMENT	and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml)		
FEATURES			
SOURCE			

complement(join(144, .2404,.2450, .4378,.4501, .4660))  
/note="ORF\_ID:dd\_02769"  
/codon\_start=1  
/product="putative protein predicted by Arabidopsis  
Challana"  
/protein\_id="AA03192.1"  
/db\_xref="GI:20976603"  
/translation="MEMMLTFEENSISIKILILPIGDISPEKRYKETSILITINIELS  
STRSNDELSPESKISWDSGLMNFVIGSGFORSSEEDLOTQKVFVGIVGWCKRS  
KLOEKFRLFEIAVAVOYPSCVSLSCDADVPDDPDGLGALMLIPMNSDRKHILFY  
LTLTIDRSHMALIKFEKMYVLSLEHCDVSTLSSQNNINNNSSGIIINSIGLGN  
MNNNSLIGSTIGGAMNSGVSSVTSNNNSLITTPLOS IKLMDEIGAKKRFGR  
LNKRCDFPELLAGSPDLAKRYDLSIDCRSNDMEWLAGSEGYISAVALIKRNOFN  
TOAGSTVPIISPNSITGSNSNSGSGNSGNIYNTSGSGNIYNTNTNSNN  
NNNNNSNTNTNNNNNNNNNNNGVGNMGNNLIDLOIGSDLKIELSEALII  
SYNKRKYPOVEYDVMLEFRANYHSMDRIEASSELLTMNAYDIDPESITDITSCIA  
LLYSXGPRFRKAFYFLREAFILYNTREPMKINHILLIASKYTOLEDLESNSRFL  
DEGSKSLSTTLNOMKSNVSYSGRNNNNNGGGGGGSGKNNNSNNSPTNVGQ  
ADYNOFPNOSIDQOVOSNSNRSSSSSSTFTNNIPELSYRKEPMGSAVLIHLYN  
LINSLEINMDSLNIISOQKOVEFOLDLHSHKALSLSINTVYHMLLGLPIVAVYS  
LOLNEPIYRPSOOSLPFAOREKNSNMFETPYPGGSGNSGIDGGIGGGGSGSN  
SNRKVMYAGGEGSVVYVTLNSPSEFPIEIOSLTLNAGVPEPESITSLTSDM  
DIVISGLESPGLIKIGVFRANULSEHPDYDYGNSISISQVNELYRIGDAYKDI  
LPEDDTTASANTSSPSSRONNNNDELIVNKAIFATPKLPIKADVPPLGVNLVLA  
SESSEPLEQNGITQIDISVLEFPAALAKRKLAANOINPELSDSEDTISFOWDSI  
ISISLPLEGFTFKLPKISFSPLEPGNOFLIDVYSGENDEEBEENNTTNTNT  
QYLTVAGNQSOTLATTTTTATYKORKLTPLSTIGHGQVINFIVASYSKISL  
SHYLVDRSSNNGLEFKVEKDODLKEDECENTYPLCITGHEDIKRNTSTHFTISSE  
LEIPINVENNKRKINNLSKSFISPGNICITIVMKRESPLDOLPLRQAKGYIKPK  
KLTIEEYVAKCLOIYKOLITSKIKLAMSNDTPNSILDNVILFSKILHLKNDQ  
VLIOERKNKOLVKTIEOGGFETIEVLVTNLNTPINCLVNLHLIPLIDQNNNNNGK  
NGNNTVDSNKLGYIGSLTTPIDLEPGATPNHSLDIVIFPERSYKFIISCELFKTKQ  
IIPASHSLVNVV"  
9249.15214, .6624, 6947, .7146, 8631, .8912, 9042, .9155,  
9424, .9457, 9627, .10028)  
/note="ORF\_ID:dd\_02770"  
/codon\_start=1  
/product="Ribonuclease inhibitor."  
/protein\_id="AA03193.1"  
/db\_xref="GI:20976604"  
/translation="MDVVOYIKDEFOETICLNDLSILEIDNODODEFISPFYK  
OSIHKIIPILHMOIIEYVYTLTGEFNSPEYTKIMDSKNDRYSLEAVCYWKSII  
KASCSILSLSNLEANTFRITHTYTKLKEITINNINDSYDEPDDSDSDSDGDN  
KRDVSNDDIYDEQKKNMTRRRKRVLNDNSDKIGCLATTIINSLSPLSTLNSIM  
TNLKRANVYIOSSLSCNRNTRKLAFOSCPIGDGVVYLSLDSNOCIRPOLDIOGSG  
VNGLEYEYGRALSNNOTLEKILMSYNSQNTSVCYASKNSGRHSQMSKLSIMGCI  
GNGALSTEDYALINKSLKEIDLGNOFGDAGATLAKLSHPSLSILDSNLSITE  
GDSILFLESQSTTTLTNSRNLDSITPLNTVTSLAENKTYLTLSNTECNOJSH  
FPOISALQSKTKIKLNTLSKHKRKSNDITPLNTYTSLAENKTYLTLSNTECNOJSH  
ITHOLEYIDLSLPHETLRNINULLREFRKILKNSPFLPEVYKAAKNKTOOQAOAISQ  
AGVLAADVLRLTJHNGDREGFRKILDAATRAKDENRDVAVALNLETFPKQALIE  
DDQKIAFEENKRRKDRQMRQDILKEATRKAKENKTRPEVLAQOITELNPKIKISE  
OLNANKLEADVIMERKHAQILSQFNOFLKEITELSDYMKRSTMLNHRDLINSO  
OLPLEBEMSIKOEERTVLOIOPGDSQADYRISYAGCTIOTGGGDYVESYEDN  
YDNYNGYDGGYGGGSELQARLALYDESAEFSDNLAKAGDIISVYIQDDDSQWTKG  
RNYBENBETIPSYTIEYV"  
complement(join(10585, .10678, 10842, .10913, 11037, .11212))  
/note="ORF\_ID:dd\_02772"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA03194.1"  
/db\_xref="GI:20976605"  
/translation="MVGTRQKWODLPDPAGGFKLKLYARTSTSPLEGATPGVCSIM  
AVGYIIFSDKVERNAREEBEERKLSMILPILOAENDINFLASPHQNVYETRWMPQOT  
GKRAAALLRD"  
complement(join(11580, .11786, 11891, .12133))  
/note="ORF\_ID:dd\_02774"  
/codon\_start=1  
/product="putative protein ubiquitin-conjugating enzyme,  
FliCIIa pastoris"  
/protein\_id="AA03195.1"  
/db\_xref="GI:20976606"  
/translation="MAARLKLKKYKVLQNEEFEDIIILPYPRDESIDLRWVALIIGPDPDT







REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
3 (bases 1 to 45713)  
Waterston,R.  
Direct Submission  
Submitted (22-MAY-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
4 (bases 1 to 45713)  
Waterston,R.  
Direct Submission  
Submitted (30-AUG-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
5 (bases 1 to 45713)  
Waterston,R.  
Direct Submission  
Submitted (27-SEP-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
6 (bases 1 to 45713)  
Waterston,R.  
Direct Submission  
Submitted (05-OCT-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
7 (bases 1 to 45713)  
Waterston,R.  
Direct Submission  
Submitted (19-APR-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

COMMENT  
Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RO, England  
email: tw@nematode.wustl.edu and jesse@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this cosmid sequence and its analysis see:  
(www.wormbase.org/db/seq/sequence?name=B0511;class=Sequence)

#### NEIGHBORING COSMID INFORMATION

The 5' cosmid is C43H8, 2850 bp overlap; the 3' cosmid is C34B2, 300 bp overlap. Actual start of this cosmid is at base position 2847 of B0511; actual end is at 45613 of B0511.

#### NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneFinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yui Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE\_INDEX.html) and The C. elegans Genome Cloning Project (http://worldb.dicf.harvard.edu/), similarity to other proteins from Blastx analyses (http://blast.wustl.edu/), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and

FEATURES	Source	misc_feature	gene	CDS
Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).				
Location/Qualifiers				
1..45713				
/organism="Caenorhabditis elegans"				
/strain="Bristol N2"				
/db_xref="taxon:6239"				
/chromosome="I"				
/clone="B0511"				
83..84				
/note="SL1 trans-splice acceptor; see yk290h4.5"				
91..2919				
/gene="B0511.6"				
/note="for a graphical representation of this gene see: (www.wormbase.org/db/seq/sequence?name=B0511.6;class=Sequence)"				
join(91..402,451..1173,1677..1979,2239..2355,2475..2576,2842..2819)				
/gene="B0511.6"				
/note="similar to DEAD-box helicases (Pfam: DEAD.hmm, score: 262.89) (Pfam: hellicase.C.hmm, score: 76.21); coded for by the following C. elegans cDNAs: cm01c4, yk2925.3, yk9295.5, CESEF01F, CESEF227F, CESEF227R, yk290h4.5, yk391e3.3, yk391e3.5, yk461a5.5, yk48201.3, yk48201.5, yk483a8.5, yk727e5.3"				
/codon_start=1				
/product="Hypothetical protein B0511.6"				
/protein_id="AAC17654.2"				
/db_xref="GI:15375231"				
/translation="MGIDYKKVLRKRLKQMDGHEKKAKRVYVEEPEDEEEVE EQKHOTEESESTKSEFLTKTFASLSEGVANLKAQVIGFTTWTETAKSDPL LEGKVLASAKTSGKTLAFLPALELHKLMKQHNGTGVIVSPRELSEQVGL SELKGSNLTGVWGSNRSAREKDLKAGVILVATGRLDHLQNTDNELVNKKC LIIDPADRLIDGFEIEMOQVLRHPKOROSMLFSAIATSPKDELVKALSNPVKRS VHEKKEAETVEGIDGGYIVAPSDRKLILFTFLKKNKKKKYVPPSSONSYSFHEHL NYIDPCKSIHKKOKQKRTTFFPQCAETGILCTVYAAAGLDIPADVNIQDPT DDEPRTIHRVGTANGTSGKALVLTPEELGFTLRYLKAAYLTNEFSMSKANT QSDNLNLSKNTYLNKSAEKAYCYLRVDSISLDIDFYVIMDLTAVSKSFGSVPP FVLDLPSNKPVKYIERSKLSCAGYRRKKOSFFEKAK"				
4800..6253				
/gene="B0511.5"				
/note="for a graphical representation of this gene see: (www.wormbase.org/db/seq/sequence?name=B0511.5;class=Sequence)"				
join(4800..4812,4910..5106,5571..5642,5691..5832,6000..6253)				
/gene="B0511.5"				
/codon_start=1				
/product="Hypothetical protein B0511.5"				
/protein_id="AAC17653.1"				
/db_xref="GI:1165535"				
/translation="MPTKGRVFLVGHSDKDCVSENGRTTSITVPDKCGVEYVGH GKAGYTSVNVIVISFHDKFTKVDRAVNTCLVAPTDVSVALTQPSLKKIOYL AOPSCVEYEDVDFRRPRAELIVHVAPELHWCTDGLDFCMTVHDCVYNEKSKSR RSKTIIDSGCSLIDTRLPNLKRENNKLSARVSKAFRRGDVAVVEFGVNLDIRNGT SCPRPRC"				
complement(6442..7843)				
/gene="B0511.7"				
/note="for a graphical representation of this gene see: (www.wormbase.org/db/seq/sequence?name=B0511.7;class=Sequence)"				
complement(join(6442..6918,7008..7178,7262..7434,7486..7585,7641..7694,7739..7843))				
/gene="B0511.7"				
/note="similar to forkhead-associated (FHA) domains (Pfam: FHA.hmm, score: 28.36); coded for by the following C. elegans cDNAs: yk393e11.3, yk465a3.3, yk465a3.5, yk626b10.5, yk718c8.3, yk718c8.5, yk727h2.5, yk728a2.3"				
/codon_start=1				
/product="Hypothetical protein B0511.7"				
/protein_id="AAC17655.1"				
/db_xref="GI:1165537"				
/translation="MEETIDSKSRKVPSPMAKPPNGAHLVDQGDALIQKLLDDKKA YFGRNNKQVDFAVEHASCVRHALLLHYGLQRPALIDMDSHGTELGNNVRLPLEY				



cdna EST EMBL:rt00171 comes from this gene"

gene  
/codon\_start=1  
/protein\_id="CA88874.1"  
/db\_xref="GI:3879998"  
/db\_xref="SPTREMBL:Q2675"  
/translation="MTIVYHOLLIVLIFIGAKGAGYVIGSGASOPNRTDVMVWVPS  
MTKNCISIDVEKGYLONEDHGVGGRKFOALFYEHSEKILPYFAONSDPKNGCIP  
OMDLELAHILQAEKDINERTIPDENFGIAVIDIEERPMMELSKGPFYKTESIRLT  
ROOHFWSTKQIEMQAEKDYERAKQAFETLIRLGRLEPNKWKYVLEPKCNGVGO  
KSDSDTSLEKFORNDNLHMLWGEESTALFYSIQKONPEYNNVSGALITTEKRI  
KRSQCSQICINIHFTKLEIENPYYPDPDFSKOMYLASTLDAIMANNSVYINSTSIS  
SNVYSQTYVDVNTIGLYQLTDRNLNDCMEGEGEGEYDLEPRTNPAIYNACRCE  
RPFEGKSCYRGRBMGVSMKASQTPQVDPYTAFFSTSSNCTKKYNAPNDQFYSNTGC  
DIKLAKKL"

gene  
complement(join(7754..7819,7868..8059,8139..8350,  
8608..8952,9038..9182))  
/gene="t22c8.3"  
complement(join(7754..7819,7868..8059,8139..8350,  
8608..8952,9038..9182))  
/gene="t22c8.3"  
/note="contains similarity to Pfam domain: PF00096 (Zinc  
finger, C2H2 type), Score=29.5, E-value=2.5e-05, N=2"  
/protein\_id="CA88875.1"  
/codon\_start=1  
/db\_xref="GI:3879999"  
/db\_xref="SPTREMBL:Q22676"  
/translation="MATTVEQLANTQYQKEPVLDLDDSNMEIAHKIQNOTKRLPISEPA  
MKNKVYSQNKETNKRTEFEKSRKATSTIDRYLRTQRTKDKESTYEKNALIKS  
AKOIVKREEEGDDLEDELTLVPSKRMWVSKNRKRTIKPNRDTIAEBQTEHGH  
VYQAFEDPQKHRCQPECLLESTPDKLRHMKSGDVPDYCSLGNPIKFKYNLRY  
HLEKCDKKKLSDHSMLSKVPDPFVKILOEAKKPVLPMTLEGMPREFLKGVYGF  
DPSLITQIMAKIYEKKORRLACIDSDSDDELDPDFDFDE"  
complement(join(9764..9814,9844..10053,10113..10324,  
10438..10872))  
/gene="t22c8.4"  
complement(join(9764..9814,9844..10053,10113..10324,  
10438..10872))  
/gene="t22c8.4"  
/note="contains similarity to Pfam domain: PF00096 (Zinc  
finger, C2H2 type), Score=19.5, E-value=0.025, N=2"  
/codon\_start=1  
/protein\_id="CA88876.1"  
/db\_xref="GI:3880000"  
/db\_xref="SPTREMBL:Q22677"  
/translation="WTVSFLEPSTFEGSLVVVIVICSTSLTLHCTAFAFKFSLKVEI  
LRHLHQTVEFLSDMNTNSTSTSGSDMDIENHSPKLSGVAPRRRIDKSTSEAK  
IORMDDAMKKLKEQEKARLKKKEETLAKSETPAAKKHMIIGAGSEDPKITALG  
DICFALFNADKKRRHDSHDEPEVQPMCNAPLTKYKYNLKKHIMCOPSEKMEG  
SMLAAYADYDLOLEAAVNPISLPMPTINGMPYMDRISPYTGDKMSLDHLDMLER  
YPTDVFQNLQISFDEADLDVNLFDER"  
join(12937..13133,14103..14201,14286..14396,14608..14701)  
/gene="t22c8.5"  
join(12937..13133,14103..14201,14286..14396,14608..14701)  
/gene="t22c8.5"  
/note="contains similarity to Pfam domain: PF00096 (Zinc  
finger, C2H2 type), Score=72.5, E-value=2.8e-18, N=3"  
/codon\_start=1  
/protein\_id="CA88877.1"  
/db\_xref="GI:3880001"  
/db\_xref="SPTREMBL:Q22678"  
/translation="MDGAKATFFRPWESHGYSIHSLDPSIPDPSPASTASSSSSSIG  
ANELTKRKKEKRCCTCPNCAIKHGDHRSQHTHLCSVGGCKTKYKSHLRAHLRKTT  
GDPRVPCWMEFGCKRFRSDQLIRHKRHTKEYEFAKCFKIPRSPRSDHLDQHLTSVH  
NIVVD"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
/codon\_start=1  
/protein\_id="CA88878.1"  
/db\_xref="GI:3880002"

gene  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"  
complement(join(15072..15356,15407..15470,15586..15728,  
15840..16043,16127..16221,16685..16796))  
/gene="t22c8.6"

gene  
CDS  
gene

```

/db_xref="SPTREMBL:Q22679"
/translation="MDREHLIKLACARNPDLEMCQOLKARNYRDSNSGKLEMLIEPPL
PSMAARRDIDIAHNDENSROKRVYRRLAEVYIILOTKCARGPVLQKHCQPKT
SARNRGAIFPRDCAQPIEKGDPLGAIASFPSGVNIIALVADYKGIIPYLIDEGE
VQVGLAIPFGWAGGASTSVGROYDFGVGANNWDKYGKKNMNIPIVQSG
VGGGQHNVSPLIGKDKAGNLKAYNGCGVGYIQQNDHGVNFKRGDVRHFFGVSPP
VAGQYQIQANAFPLGDVYBERALG"
complement(join(17992..18093,18136..18440,18799..19024,
19407..19781,19822..19987,20032..20250,20393..20535,
20583..20679,20990..21078,21416..21523))
/gene="T2C8.7"
complement(join(17992..18093,18136..18440,18799..19024,
19407..19781,19822..19987,20032..20250,20393..20535,
20583..20679,20990..21078,21416..21523))
/gene="T2C8.7"
/note="N-terminal portion of this predicted gene show
strong similarity to C. elegans cuticulin"
/codon_start=1
/protein_id="CAA88879.1"
/db_xref="GI:3880003"
/db_xref="SPTREMBL:Q22680"
/translation="MPFVNCFMILFFVKNVIANDSNNFYTHMEPPKLECGSEGIR
LHNPDTGGEGHVYVGEFPGQVCHLCTCLNPIYMDLPFGPCVRRRRVAPR
SISDVYIIQHPHLPISPDKAYRLNCTIYROESTLQORINVSIDPSTALSNAPK
CRVYLSGLNGPYRPAANGVDVYHAKMTCDSDFRVHSCVYRDESGKQFQIDR
GCVDITLSPFVYSDDLKSAFTVAARAFYAAQVAVHSCQITTCQKQENCCBISPP
ICRPMDDGLPIKIVHYKHKNGKEGFDGGEVETLPRENTSGSTTVYRSTTTERAO
AFSTTKTKTTSSSLQTLKYQOREDDYSLNESLSELSGFEVORLSDETRDKLTKR
TSPILRNNSSPSSKRRNGYNPNVESEGVDGONTPQVYKRKSSDAGVNEGIGI
AKDSQAKOKKISLVAPREKLPKRVFHEHTISKMDSELSPIPKTASYCHORSR
SHKVEIDANTILEVETKTMILENPSAKERLRLEADGSGPPEPMKCTTSNTVNT
LLTLTGLQVLTSLALITQRFYKSIDNLIYADDRGSRKF"
complement(19166..19249)
/gene="T2C8.t1"
complement(19166..19249)
/gene="T2C8.t1"
/product="tRNA-Leu"
/note="TAA Leu tRNA
predicted using tRNAscan-SE-1.11
```

Query Match	Similarity	65.1%	Pred. No. 1.9	Score 46.8	DB 3	Length 27770
Best Local	69	Conservative	0	Mismatches	37	Indels 0; Gaps 0;
QY	451	CCAGTATGCCGTTATGAATTTTGGATGTGGTGACCAACCGTCACACGATTCATTGCT	510			
Db	20213	CCAAATATGCGATATGATGATGACTCTCCGGTTCACATAATGTGCACAGTTGTCATTGCT	20154			
QY	511	ATCATGTGTCAGCAGCTTATATCATTAATGACATGCGATTCGATAA	556			
Db	20153	AATGTGGAGATGTACTGTCTTACACAGTGCACATGTGATTCAGATA	20108			
RESULT 38	AF455359	6942 bp	DNA	linear	BCT 11-FEB-2002	
LOCUS	AF455359					
DEFINITION	Spiroplasma citri GI13 cell shape-determining gene cluster, complete sequence.					
ACCESSION	AF455359					
VERSION	AF455359.1	GI:17981532				
KEYWORDS						
SOURCE						
ORGANISM	Spiroplasma citri. Spiroplasma citri Bacteria; Firmicutes; Mollicutes; Entomoplasmatales; Spiroplasmataceae; Spiroplasma. 1 (bases 1 to 6942) Machetoni, W. Jr., Joel, R., Bove, J.M. and Garnier, M. Identification, organization and expression of five mreb-1-like genes in Spiroplasma citri: Effect of mreb1 disruption in insect transmission and phytopathogenicity unpublished 2 (bases 1 to 6942) Machetoni, W. Jr., Joel, R., Bove, J.M. and Garnier, M.					
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
AUTHORS						

TITLE	Direct Submission
JOURNAL	Submitted (04-DEC-2001) Genetics, ESALQ/USP, Av. Carlos Botelho 11, Piracicaba, SP 13400-970, Brazil
FEATURES	location/Qualifiers
source	1. .6942

BASE COUNT	2641 a	879 c	1096 g	2326 t
ORIGIN	YHNERSQOITGRDIVSGLPREAKISSEIRNVLNAFSKITDLYIELLENTPELACD IMRGTIVCGGALIRINDKYFDFIQLPKIASDSANCVIEGKIIEKTIKKNIENG LYNFQEGKLLSTLGKKRK"			

gene	1076.	.2125
CDS	1076.	.2125

Query Match	3.9%	Score 44.8	DB 1	Length 6942
Best local Similarity	44.6%	Pred. No. 5.8		
Matches 175	Conservative 0	Mismatches 217	Indels 0	Gaps 0
OY	195	TGATGAGGTGGACCTCAAGTTGCCGGAAATTTCACTTCATTGATTTATGCAATGTTC	254	

```

/note="actin-like protein
/codon_start=1
/rnastl_table=4
/product="MreB3"
/protein_id="AAL51057.1"
/db_xref="GI:17981533"

```

[illegible]

NSLITLGCNEAVKMGKTHDIRMVT  
SVLILACPSVTELRGALKATAKMD  
IGCGTTDIAIISAGDIVISKSVKVAO  
ALVKINERTVAFGRDITITPRELAI  
LAGDIRNGITICGGGALLRGIVKYE  
IERIDLPAANEKIS"

[illegible]

```

/feature="mre3"
2385..2443
/feature="mre3"
/note="actin-like protein"
/codon_start=1
/transl_table=4

```

QY 435 TCAATTTCTCCGATGCCAGATGCCATTATGAATTTTGGATGCTGGACCAACCGGACA 494

Db 6279 TGCCTTTCTAAATTTACAGATTAGTAATTGAATTAATTAGAAATTTACACCAACCGAATT 6338

QY 495 ACCAGTTCAATTTGCTATCATTTGGTCACCAAGTTATATATAAATGGACATGCGATTCTGA 554

```

/protein_id="AAL51058.1"
/db_xref="GI:17981534"
/translation="MAISDLKNTFN
ASVAYEIGTKLIAGNDARKLIG
AKVADIMKNATVLIACPKNVPLEK
APKGTIILDIGGKSSCAITISAGIV

```

OY		555	AACCGTGTACTCTTCGCGGTTGCCATT	586
Dd	6399	TATCATTAATATTTCTTGATATTTCCAA	6430	

```

EVLEKTPAEILAGDAVANGILTVGGCC
AYEKKNRDILIEKKNR"
4424..5467
/gene="mreB"
4424..5467
CDS
/gene="mreB"

```

AL772373/c	229420 bp	DNA	linear	HTG 17-JUL-2002
LOCUS	AL772373			
DEFINITION	Dado rerio clone CH21-143A2, *** SEQUENCING IN PROGRESS ***, 23 unordered pieces.			
ACCESSION	AL772373			
VERSION	AL772373.5	GI:21911694		

```

/codon_start=1
/translation_table=4
/product="Mreb4"
/protein_id="AAL51059.1"
/db_xref="GI:17981535"
/translation="MALFNSAKKPAF

```

SOURCE	ORGANISM
Danio rerio.	Danio rerio.
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 229420)	

gene  
5521..6579

**TITLE** Direct Submission  
**JOURNAL** Submitted (16-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk  
**COMMENT** Clone requests: clonerequest@sanger.ac.uk  
On Jul 19, 2002 this sequence version replaced gi:21738685.  
----- genome Center

```

CDS      5521..6579
         /gene="mreB5"
         /note="actin-like protein"
         /codon_start=1
         /transl_table=4
         /product="MreB5"

```

```

Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zface@sanger.ac.uk
-----
Project Information
Center project name: ZC143A22
-----
Summary Statistics

```

```

/db.xref="GI:17981536"
/translation="MRPETRPISLDGTAVLAVVSGQGVVNEPSLMAVNNKTNISI
IALCKAAYDAVKGKTHGDIRVTPPLYDGYIADMEAAQDLKRI FSKMRKMTNKA
IVTLACSGVTELEERALKVNAKMGAEIYVIEEPAKMAAGAAINELPGHILIDGGG
YTDLAIISSGDIIVRSRIKVAGNHDDDIRKYINSEYIAIGQKAEDEVKKFISGLV

```

Chemistry: Dye-terminator;100% of reads  
Consensus quality: 218244 bases at least Q40  
Consensus quality: 221659 bases at least Q30  
Consensus quality: 224352 bases at least Q20  
Insert size: 227220; sum-of-contigs



RESULT	40
AL645495/c	
LOCUS	81091 bp DNA linear PRI 28-NOV-2001
DEFINITION	Human DNA sequence from clone RP13-472L18 on chromosome 13, complete sequence.
ACCESSION	AL645495
VERSION	AL645495.4 GI:17154392
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE 1 (bases 1 to 81091)

**TITLE** Direct Submission  
**JOURNAL** Submitted (28-NOV-2001) Wellcome Trust Sanger Institute, Hinxton

**COMMENT**

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr13>

RP13-472L18 is from the library RPci-13.2 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP13-472L18. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP13-472L18 is at 81091 in this sequence. The true right end of clone RP11-427O13 is at 2000 in this sequence.

FEATURES	Location/Qualifiers
source	1. .81091

BASE COUNT	26912	a	14138	c	14152	g	25889	t
ORIGIN								

Query Match	3.7%	Score 43;	DB 9;	Length 81091;
Best Local Similarity	50.2%;	Pred. NO. 15;		
Matches 106;	Conservative 0;	Mismatches 105;	Indels 0;	Gaps 0;

OY	192	TAAATCATTAAGGTGGACGCTCAAGTTCGCCGAATTTCACCTCCATTTGANTTCATGAAGT	251
Db	51904	TAAAGTTCTCTCTTATCTGAGGAGCGCAGATATGCTAGTTTTTAGCTATTAGCCAAAGT	51845
OY	252	TGCAGTACAGATCTCGAATCCAGCGTGATTTTTTGAACAACAACACTGTTGTCATTTC	311

Db 51844 GAGGGAACAGAGATCTAAGATGCGCCTTTTAAGTTGAATATATATGCTGGGACGAGCC 51785

QY 312 GTTTCATCCATTTATTTGTTTCAACAAAGTGATGCGTCATATGAGACAAAGCTTTTCAAT 371

Db 51784 TCACCATTCATGCTTTGGTACTGCGCTAACGCTGTTTCACTAAAGCCTAGATTTTATATGG 51725

QY 372 GGAAGCTGATAAACAGTTAGTCGACAGATT 402

Db 51724 TTTAAGGAGCTTAAAGACTTTTGTGTGAGAGCTT 51594

```
Search completed: March 28, 2003, 16:44:58
Job time : 4054 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 13:31:58 ; Search time 304 Seconds

(without alignments)  
8600.553 Million cell updates/sec

Title: US-10-054-562a-3

Perfect score: 1161  
Sequence: 1 atgatgattcgtctattgc.....ttcgtccaatcagaagca 1161

Scoring table: IDENTITY\_NDC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1161	100.0	1161	22	AAH24828	Nucleotide sequenc
2	1161	100.0	1161	22	AAH24829	Nucleotide sequenc
3	1161	100.0	1161	22	AAH24826	Nucleotide sequenc
4	1161	100.0	1161	22	AAH24827	Nucleotide sequenc
5	1161	100.0	1161	22	ABL57372	Dirofilaria immiti
6	504.2	43.4	1372	22	AAH24830	Nucleotide sequenc
7	504.2	43.4	1372	22	AAH24831	Nucleotide sequenc
8	504.2	43.4	1372	24	ABL57373	Dirofilaria immiti
9	461	39.7	892	22	AAH24839	Nucleotide sequenc

C	10	461	39.7	892	22	AAH24840	Nucleotide sequenc
C	11	461	39.7	892	24	ABL57374	Brugia malayi cuti
C	12	310.4	26.7	813	22	AAH24832	Nucleotide sequenc
C	13	310.4	26.7	813	22	AAH24833	Nucleotide sequenc
C	14	40	3.4	1664976	19	AAV21209	Methanococcus jann
C	15	39.4	3.4	9810	24	ABL32426	Human immune syste
C	16	38.8	3.3	14066	20	AAH24835	Nucleic acid sequenc
C	17	38	3.2	6799	24	ABK39779	CDNA encoding full
C	18	37.2	3.2	6242	24	ABO70961	Listeria monocytog
C	19	37	3.2	783	20	AAZ16979	Human gene express
C	20	37	3.2	783	20	AAH99054	Human gene express
C	21	37	3.2	795	20	AAZ17515	Human gene express
C	22	37	3.2	795	20	AAH98778	Human gene express
C	23	37	3.2	1467	21	AAH29937	Human secreted pro
C	24	37	3.2	1467	22	AAH33034	Human colon cancer
C	25	37	3.2	1680	22	ABO70472	Listeria monocytog
C	26	36.8	3.2	6571	22	AAH89354	Arabidopsis thalia
C	27	36.6	3.2	1527	22	AAH65681	Nucleotide sequenc
C	28	36.6	3.2	8342	24	ABL32500	Human immune syste
C	29	36.4	3.1	1163020	24	ABO67197	Listeria innocua C
C	30	36.4	3.1	3011208	24	ABO69245	Listeria innocua D
C	31	36.2	3.1	510	22	ABA84521	Human breast cell
C	32	36.2	3.1	510	22	ABA84521	Human breast cell
C	33	36.2	3.1	510	22	ABA84521	Human foetal liver
C	34	36.2	3.1	510	22	ABA84521	Probe #11883 for g
C	35	36.2	3.1	510	22	AAK14772	Human brain expres
C	36	36.2	3.1	510	22	AAK40512	Human bone marrow
C	37	36.2	3.1	1103	24	ABK77860	Probe #15232 used
C	38	36.2	3.1	7810	22	AAH45434	Bacillus clausii g
C	39	36.2	3.1	7810	22	AAH45434	Chemically pretrea
C	40	36.2	3.1	1481	22	AAH28281	DNA transcription
C	41	36	3.1	1481	22	AAH43068	Human colon cancer
C	42	35.8	3.1	44362	23	ABL29558	Human ovarian anti
C	43	35.6	3.1	541	18	AAH83869	Drosophila melanog
C	44	35.6	3.1	1195	24	ABO68674	DNA encoding a N-a
C	45	35.6	3.1	2365589	24	ABA90521	Listeria monocytog

#### ALIGNMENTS

RESULT 1	AAH24828	standard; DNA; 1161 BP.
ID	AAH24828	
XX	AAH24828;	
AC	22-AUG-2001	(first entry)
DT		
XX		
DE		Nucleotide sequence of a cuticlin polypeptide.
XX		
KW		Cuticlin; gene therapy; vaccine; helminth parasite; ss.
XX		
OS		Dirofilaria immitis.
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1161
FT		/*tag= a
FT		/product= "cuticlin"
FT		/note= "no termination codon given"
XX		
PN	US6248329-B1.	
XX		
PD	19-JUN-2001.	
XX		
PF	01-JUN-1999;	99US-0323427.
XX		
PR	01-JUN-1998;	98US-0087435.
XX		
PA	(CHAN/) CHANDRASHEKAR R.	
XX	(MORA/) MORALES T H.	
PI	Chandrashekar R, Morales TH;	

xx	MPI: 2001-396953/42.
DR	P-PSDB: AAB84258.
xx	
PT	Novel nucleic acid molecule encoding a <i>Dirofilaria immitis</i> cuticlin
PT	protein is useful as a vaccine to prevent parasitic helminth infection
PT	-
xx	
PS	Claim 1; Column 35-40; 29pp; English.
CC	The specification describes a <i>Dirofilaria immitis</i> nucleic acid molecule,
CC	that encodes a cuticlin protein. The <i>Dirofilaria immitis</i> nucleic acid
CC	molecule is useful as a probe to identify nucleic acid molecules, as a
CC	primer to produce nucleic acids, as a therapeutic reagent, e.g., gene
CC	therapy to inhibit cuticlin activity or production, or in a vaccine to
CC	prevent infection with helminth parasites. The cuticlin protein,
CC	antibodies raised against it, and inhibitory compounds of cuticlin may
CC	all be used in compositions to protect animals, especially mammals such
CC	as cats, dogs, and humans. The antibodies may be used to passively
CC	immunize an animal, or as reagents in assay to detect infection of
CC	helminths, or as tools to screen expression libraries to recover desired
CC	proteins. They may also be used to target cytotoxic agents to the
CC	parasite and kill it directly. The present sequence encodes a cuticlin
CC	polypeptide.
xx	
SQ	Sequence 1161 BP, 338 A; 227 C; 235 G; 361 T; 0 other;
	Query Match            100.0%; Score 1161; DB 22; Length 1161;
	Best Local Similarity   100.0%; Pred. No. 0;
	Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ATGATGATTGCTTTTATTCCTTTCGTGTAACACTTATGCATTTGCTTATTCGATTCGG 60
Db	1 ATGATGATTTCGCTTATTCCTTTCGTGTAACACTTATGCATTTGCTTATTCGATTCGG 60
OY	61 GTTGCAATGGTGTGGAGGTGAGCCAGAAATGGAATGTGGACCACCTCATTAACAATC 120
Db	61 GTTGCAATGGTGTGGAGGTGAGCCAGAAATGGAATGTGGACCACCTCATTAACAATC 120
OY	121 AATTTTAATACACGTAAATGATTCGAGAAGACATGTTTATGAAAAGTCTTATATGATCAA 180
Db	121 AATTTTAATACACGTAAATGATTCGAGAAGACATGTTTATGAAAAGTCTTATATGATCAA 180
OY	181 GAAGSTTGCCGTAAATGATGAGAGGTGAGCGTCAAGTTCGCCGAATTTCACTTCATTTGAT 240
Db	181 GAAGSTTGCCGTAAATGATGAGAGGTGAGCGTCAAGTTCGCCGAATTTCACTTCATTTGAT 240
OY	241 TCATCAGATGTTGGCGGTACACGATCTCGAATCCACGATGAGATTTTGTAAACAACAAT 300
Db	241 TCATCAGATGTTGGCGGTACACGATCTCGAATCCACGATGAGATTTTGTAAACAACAAT 300
OY	301 GTTGCATTTGCTTTCATCATCATTTATTTGTACCAAGTTGATCGTGCAATGAGTAGACA 360
Db	301 GTTGCATTTGCTTTCATCATCATTTATTTGTACCAAGTTGATCGTGCAATGAGTAGACA 360
OY	361 TGCTTTTACATGAGAGCTGATAAACAAGTATAGTGACACAGATTGAGATATCGAAATCACA 420
Db	361 TGCTTTTACATGAGAGCTGATAAACAAGTATAGTGACACAGATTGAGATATCGAAATCACA 420
OY	421 ACTGTTTTCAACATCAATTTGTCCGATGCCAGTATGCCGTTATGAATTTTGGATGGT 480
Db	421 ACTGTTTTCAACATCAATTTGTCCGATGCCAGTATGCCGTTATGAATTTTGGATGGT 480
OY	481 GGACCAACCCGGTCAACAGTTCATATTTGCTATCATTTGGTCAAGCCAGTTATCATATAAAGG 540
Db	481 GGACCAACCCGGTCAACAGTTCATATTTGCTATCATTTGGTCAAGCCAGTTATCATATAAAGG 540
OY	541 ACATGCGATTCGTGAACCCGTTGATACTTTCGCGCGGTGTGCCATTCCTGCTTGTGCAT 600
Db	541 ACATGCGATTCGTGAACCCGTTGATCTTTCGCGCGGTGTGCCATTCCTGCTTGTGCAT 600
OY	601 GATGCTAACGGTGAATCTGTGGAATTTCTAATGCTGATGATGTGCTTGTGATTAATAT 660
Db	

Db	601	GATGGTAAACGGTGAATCTGTGGAAATTCTTAAATGCTGATGATGATGCTCTTGAATAAAT	660
Qy	661	TTGCTCAATTAATTTGGATATATCCAAAGATTTTAATGGCTGGCCAAAGATCAGCTATATC	720
Db	661	TTGCTAAATTAATTTGGATATATCCAAAGATTTTAATGGCTGGCCAAAGATCAGCTATATC	720
Qy	721	AAATATGGGATCGATCAGACCTTTTCTATCAATGCCAGATCAGATTAATACATTTAAAGAA	780
Db	721	AAATATGGGATCGATCAGACCTTTTCTATCAATGCCAGATCAGATTAATACATTTAAAGAA	780
Qy	781	CCAAATPACGSAATGGTTGCGACCAATGTTCCAGAACCCAAAGAGATTTGGAGCTGTAA	840
Db	781	CCAAATPACGSAATGGTTGCGACCAATGTTCCAGAACCCAAAGAGATTTGGAGCTGTAA	840
Qy	841	ACAGGTGGTGGCGGAGCAAAACCTGCTGCAGCTGGCCAACTTCTTACTCAAGAAAAA	900
Db	841	ACAGGTGGTGGCGGAGCAAAACCTGCTGCAGCTGGCCAACTTCTTACTCAAGAAAAA	900
Qy	901	TCGTCAGAACCGGAGAAATATCATTGATGTCAGAACTGATATACACACCCTTGAATATAGC	960
Db	901	TCGTCAGAACCGGAGAAATATCATTGATGTCAGAACTGATATACACACCCTTGAATATAGC	960
Qy	961	GATGATATATCAAGCTTTGGCAGTTGATTTACGTACACCGTGCACTTCTGCAACATATATGA	1020
Db	961	GATGATATATCAAGCTTTGGCAGTTGATTTACGTACACCGTGCACTTCTGCAACATATATGA	1020
Qy	1021	CAACTGTAAATACCTTGGTCGAGTACAAATAAGSAAATCGCATGTCACCACTTGGCTTCA	1080
Db	1021	CAACTGTAAATACCTTGGTCGAGTACAAATAAGSAAATCGCATGTCACCACTTGGCTTCA	1080
Qy	1081	ATGTTTATGGGTTTAAAGCATTTGCAATTTGCTGCGCTGATTAATTAACATTTTCGTTTAA	1140
Db	1081	ATGTTTATGGGTTTAAAGCATTTGCAATTTGCTGCGCTGATTAATTAACATTTTCGTTTAA	1140
Qy	1141	TTTCGTCCAATTCAGAAAGCA	1161
Db	1141	TTTCGTCCAATTCAGAAAGCA	1161
RESULT 2			
AAH24829/c			
ID	AAH24829 standard; DNA; 1161 BP.		
XX			
AC	AAH24829;		
XX			
DT	22-AUG-2001 (first entry)		
XX			
DE	Nucleotide sequence of antisense strand of cuticlin open reading frame		
XX			
KW	Cuticlin; gene therapy; vaccine; helminth parasite; ss.		
XX			
OS	Dirofilaria immitis.		
XX			
PN	US6248329-B1.		
PD	19-JUN-2001.		
XX			
PE	01-JUN-1999; 99US-0323427.		
XX			
PR	01-JUN-1998; 98US-0087435.		
XX			
PA	(CHAN/) CHANDRASHEKAR R.		
XX	(MORA/) MORALES T H.		
PI	Chandrashekar R, Morales TH;		
XX			
DR	WPI; 2001-396953/42.		
XX			
PT	Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin		
PT	protein is useful as a vaccine to prevent parasitic helminth infection		
PS	Claim 1; Column 41-42; 29pp; English.		

XX The specification describes a *Dirofilaria immitis* nucleic acid molecule, CC that encodes a cuticlin protein. The *Dirofilaria immitis* nucleic acid CC molecule is useful as a probe to identify nucleic acid molecules, as a CC primer to produce nucleic acids, as a therapeutic reagent, e.g., gene CC therapy to inhibit cuticlin activity or production, or in a vaccine to CC prevent infection with helminth parasites. The cuticlin protein, CC antibodies raised against it, and inhibitory compounds of cuticlin may CC all be used in compositions to protect animals, especially mammals such CC as cats, dogs, and humans. The antibodies may be used to passively CC immunize an animal, or as reagents in assay to detect infection of CC helminths, or as tools to screen expression libraries to recover desired CC proteins. They may also be used to target cytotoxic agents to the CC parasite and kill it directly. The present sequence is the complement CC of an open reading frame encoding a cuticlin polypeptide.

XX Sequence 1161 BP; 361 A; 235 C; 227 G; 338 T; 0 other:

Query Match 100.0%; Score 1161; DB 22; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATATGATTCGCTTATGCTTCTGCTACACTTATTCGATTCGTTATGATCCG 60  
DB 1161 ATGATGATTCGCTTATGCTTCTGCTACACTTATTCGATTCGTTATGATCCG 1102  
QY 61 GTTGACATGCTGTCGAGAGTGAGCCGAAATGAATGAGACCACTCAATACATC 120  
DB 1101 GTTGACATGCTGTCGAGAGTGAGCCGAAATGAATGAGACCACTCAATACATC 1042  
QY 121 AATTTAATACGTAATGATGATTCGAAAGACATGTTTATGAAAGCTTTATGATCAA 180  
DB 1041 AATTTAATACGTAATGATGATTCGAAAGACATGTTTATGAAAGCTTTATGATCAA 982  
QY 181 GAAGGTTGCCGTAATGATGAGGTGACGTCGAAAGTCCGGAATTCACATTTGAT 240  
DB 981 GAAGGTTGCCGTAATGATGAGGTGACGTCGAAAGTCCGGAATTCACATTTGAT 922  
QY 241 TCATGAATGTTGCGGTACAGATCCTGAATCCAGTGTATTTTGAACAACACAT 300  
DB 921 TCATGAATGTTGCGGTACAGATCCTGAATCCAGTGTATTTTGAACAACACAT 862  
QY 301 GTTGCAATTCGTTTCCATTCATTTGTTACCAAAAGTTGATGTCGATTCGATCAA 360  
DB 861 GTTGCAATTCGTTTCCATTCATTTGTTACCAAAAGTTGATGTCGATTCGATCAA 802  
QY 361 TCGTTTACATGAGAGCTGATTAACAGATTAGTCACAGATTGAGTATCGAATCACA 420  
DB 801 TCGTTTACATGAGAGCTGATTAACAGATTAGTCACAGATTGAGTATCGAATCACA 742  
QY 421 ACATGCTTTCAAACTCAATTTGTCGAGTCCGATGCGGTATGAAATTTGGAGGT 480  
DB 741 ACATGCTTTCAAACTCAATTTGTCGAGTCCGATGCGGTATGAAATTTGGAGGT 682  
QY 481 GGACCAACCGGTACACAGTTCAATTTGCTATGTTGTCAGCCAGTTTATCAATAATG 540  
DB 681 GGACCAACCGGTACACAGTTCAATTTGCTATGTTGTCAGCCAGTTTATCAATAATG 622  
QY 541 ACATGCGATTCTGAAACCGTTGATTTCTGCGCGGTGTCATTCCTGCTTTGTCAT 600  
DB 621 ACATGCGATTCTGAAACCGTTGATTTCTGCGCGGTGTCATTCCTGCTTTGTCAT 562  
QY 601 GATGCTACCGGTGATCTGTTGAATTTCTAATGCTGATGAGTGTCTTTGATAAATAT 660  
DB 561 GATGCTACCGGTGATCTGTTGAATTTCTAATGCTGATGAGTGTCTTTGATAAATAT 502  
QY 661 TTGCTAAATATTTGGAATTTCAACAGATTATGCTGCGCAAGAGCTACGTTATAC 720  
DB 501 TTGCTAAATATTTGGAATTTCAACAGATTATGCTGCGCAAGAGCTACGTTATAC 442  
QY 721 AATATGCGGATCGATCAGCTTTTCTATCAATGCCAGATAGTATACATTAAGAA 780  
DB 441 AATATGCGGATCGATCAGCTTTTCTATCAATGCCAGATAGTATACATTAAGAA 382

QY 781 CCAATAGCGAATGCTTTCGACCACAATGTTCAAGAACACAGAGATTGAGAGCTTTAAA 840  
DB 381 CCAATAGCGAATGCTTTCGACCACAATGTTCAAGAACACAGAGATTGAGAGCTTTAAA 322  
QY 841 ACAGGTGTCGCGAGCAAAACCTGTCGAGCTGCGCACTGCTTACTCAAGAAAAGA 900  
DB 321 ACAGGTGTCGCGAGCAAAACCTGTCGAGCTGCGCACTGCTTACTCAAGAAAAGA 262  
QY 901 TCTGCAACACCGGAGAAATATCATTTGATGTCGAACTATATCAACACCTTTGAATAGC 960  
DB 261 TCTGCAACACCGGAGAAATATCATTTGATGTCGAACTATATCAACACCTTTGAATAGC 202  
QY 961 GATGATATCAAGCTTTGCGCAGTTGATTTAGTCACCGTCACTTTGCAACATATGA 1020  
DB 201 GATGATATCAAGCTTTGCGCAGTTGATTTAGTCACCGTCACTTTGCAACATATGA 142  
QY 1021 CAACCTGTAATCTTGTGTCAGTACAAATGAAATGCAATGTCACCAATTTGGCTTCA 1080  
DB 141 CAACCTGTAATCTTGTGTCAGTACAAATGAAATGCAATGTCACCAATTTGGCTTCA 82  
QY 1081 ATGTTTATGCTTAAACATTCATTCATTCATTCGTCGCTCATTTATACATTTGCTTAAA 1140  
DB 81 ATGTTTATGCTTAAACATTCATTCATTCATTCGTCGCTCATTTATACATTTGCTTAAA 22  
QY 1141 TTTCGTCCAATCAGAGCA 1161  
DB 21 TTTCGTCCAATCAGAGCA 1

RESULT 3  
AAH24826  
ID AAH24826 standard; DNA; 1779 BP.  
XX  
AC AAH24826;  
XX  
DT 22-AUG-2001 (first entry)  
XX  
DE Nucleotide sequence of sense strand of cuticlin gene.  
XX  
KW Cuticlin; gene therapy; vaccine; helminth parasite; ss.  
XX  
OS *Dirofilaria immitis*.  
XX  
FH Key location/Qualifiers  
FT CDS 167..1330  
FT /tag= a  
FT /product= "cuticlin"  
XX  
PN US6248329-B1.  
PD 19-JUN-2001.  
XX  
PF 01-JUN-1999; 99US-0323427.  
XX  
PR 01-JUN-1998; 98US-0087435.  
XX  
PA (CHAN/) CHANDRASHEKAR R.  
PA (MORA/) MORALES T H.  
PI Chandrashekar R, Morales TH;  
XX  
XX WPI; 2001-396953/42.  
DR P-PDB; AAB84258.  
XX  
XX Novel nucleic acid molecule encoding a *Dirofilaria immitis* cuticlin  
PT protein is useful as a vaccine to prevent parasitic helminth infection  
XX  
XX Claim 1; Column 33-34; 29pp; English.  
XX  
XX The specification describes a *Dirofilaria immitis* nucleic acid molecule,  
CC that encodes a cuticlin protein. The *Dirofilaria immitis* nucleic acid

molecule is useful as a probe to identify nucleic acid molecules, as a primer to produce nucleic acids, as a therapeutic reagent, e.g., gene therapy to inhibit cuticlin activity or production, or in a vaccine to prevent infection with helminth parasites. The cuticlin protein, antibodies raised against it, and inhibitory compounds of cuticlin may all be used in compositions to protect animals, especially mammals such as cats, dogs, and humans. The antibodies may be used to passively immunize an animal, or as reagents in assay to detect infection of helminths, or as tools to screen expression libraries to recover desired proteins. They may also be used to target cytotoxic agents to the parasite and kill it directly. The present sequence encodes a cuticlin polypeptide.

Sequence 1779 BP; 575 A; 327 C; 302 G; 575 T; 0 other;

Query Match 100.0%; Score 1161; DB 22; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATGATTCGTCCTTAATTCCTTCTGACTACACTTATGATTCGTTATTCGATCCG 60
DB 167 ATGATGATTCGTCCTTAATTCCTTCTGACTACACTTATGATTCGTTATTCGATCCG 226
QY 61 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATTGAATGTGGACCACTCAATACATC 120
DB 227 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATTGAATGTGGACCACTCAATACATC 286
QY 121 AATTTAATACAGTAAATGATTCGAAAGGAGCATGTTATGTGAAGGTCCTTATGATCA 180
DB 287 AATTTAATACAGTAAATGATTCGAAAGGAGCATGTTATGTGAAGGTCCTTATGATCA 346
QY 181 GAAGGTGCGGTATGATGATGAAGGTGAGCTCAAGTTGCCGGAATTCACCTCCATTTGAT 240
DB 347 GAAGGTGCGGTATGATGATGAAGGTGAGCTCAAGTTGCCGGAATTCACCTCCATTTGAT 406
QY 241 TCATGCAATGTGGCGGTACAGATCTCGAATCCACGCTGATTTTGTGAACAACT 300
DB 407 TCATGCAATGTGGCGGTACAGATCTCGAATCCACGCTGATTTTGTGAACAACT 466
QY 301 GTTGATCATTCGTTTCATCATTAATTTGTACCAAGTTGATCGTCATATGAGATACAA 360
DB 467 GTTGATCATTCGTTTCATCATTAATTTGTACCAAGTTGATCGTCATATGAGATACAA 526
QY 361 TCGTTTACATGAGAGCTGATAAACAGTATGTCACAGATTCGAGGTATCTGAATACACA 420
DB 527 TCGTTTACATGAGAGCTGATAAACAGTATGTCACAGATTCGAGGTATCTGAATACACA 586
QY 421 ACTGCTTTCAACCAATTTGTCGCGATGCCAGTATGCCGTTATGAATTTTGGATGCT 480
DB 587 ACTGCTTTCAACCAATTTGTCGCGATGCCAGTATGCCGTTATGAATTTTGGATGCT 646
QY 481 GGACCAACCGGTCAACCAATTTGTCATCATTCGTCACAGTTATATCAATAATGG 540
DB 647 GGACCAACCGGTCAACCAATTTGTCATCATTCGTCACAGTTATATCAATAATGG 706
QY 541 ACATGCGATTTCGAAACCGTTGATCTTCGCGGCTGTCATTCCTGCTTGTGAT 600
DB 707 ACATGCGATTTCGAAACCGTTGATCTTCGCGGCTGTCATTCCTGCTTGTGAT 766
QY 601 GATGGTAACGGTATACGTGTGAATTTCTAATGCTGATGAGATGTGCTTTATTAATAT 660
DB 767 GATGGTAACGGTATACGTGTGAATTTCTAATGCTGATGAGATGTGCTTTATTAATAT 826
QY 661 TTGCTAAATTAATTTGGAATATCCACAGATTTAATGGCTGGCCAGAAAGCTCAGATATC 720
DB 827 TTGCTAAATTAATTTGGAATATCCACAGATTTAATGGCTGGCCAGAAAGCTCAGATATC 886
QY 721 AATATGCGGATGATCAACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAGAA 780
DB 887 AATATGCGGATGATCAACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAGAA 946
QY 781 CCAATATGCGGATGATCAACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAGAA 840

```

```

DB 947 CCAATATGCGGATGATCAACAGCTTTGACCAACCAATGTCAGAACCAAGATTCGAGCTGTAA 1006
QY 841 ACAGGTGGTGGCCGACCAACCAATGTCAGAGTGGCCAACTGCTTACTCAAGAAAGA 900
DB 1007 ACAGGTGGTGGCCGACCAACCAATGTCAGAGTGGCCAACTGCTTACTCAAGAAAGA 1066
QY 901 TCTGCAACCGGAGAAATATCATTTGATTCAGCAAGCTGATATCAACACCTTGAATTAAGC 960
DB 1067 TCTGCAACCGGAGAAATATCATTTGATTCAGCAAGCTGATATCAACACCTTGAATTAAGC 1126
QY 961 GATGATATCAACGCTTGGCCAGTTGATTAAGTCAACCGTGCACCTCTGCAACATATGA 1020
DB 1127 GATGATATCAACGCTTGGCCAGTTGATTAAGTCAACCGTGCACCTCTGCAACATATGA 1186
QY 1021 CAACCGTATATCTGCTGCTGATACCAAAATGGAATCTGATGTCACCATTTGGCTTCGA 1080
DB 1187 CAACCGTATATCTGCTGCTGATACCAAAATGGAATCTGATGTCACCATTTGGCTTCGA 1246
QY 1081 ATGTTTATGAGTTTAAAGCTTGCATTTGATTCGCTGCTGATTTACCATTTGCTTTAA 1140
DB 1247 ATGTTTATGAGTTTAAAGCTTGCATTTGATTTGATTCATTTGCTGCTGATTTACCATTTGCTTTAA 1306
QY 1141 TTTGCTCAATCAGAGGCA 1161
DB 1307 TTTGCTCAATCAGAGGCA 1327

```

#### RESULT 4

AAH24827/C  
ID AAH24827 standard; DNA; 1779 BP.

AAH24827;

22-AUG-2001 (first entry)

Nucleotide sequence of antisense strand of cuticlin gene.

Cuticlin; gene therapy; vaccine; helminth parasite; ss.

Dirofilaria immitis.

US648329-B1.

19-JUN-2001.

01-JUN-1999; 99US-0323427.

01-JUN-1998; 98US-0087435.

(CHAN/) CHANDRASHEKAR R.

(MORA/) MORALES T H.

Chandrashekar R, Morales TH;

WPI; 2001-396953/42.

Novel nucleic acid molecule encoding a *Dirofilaria immitis* cuticlin protein is useful as a vaccine to prevent parasitic helminth infection

Claim 1; Column 35-36; 29pp; English.

The specification describes a *Dirofilaria immitis* nucleic acid molecule, that encodes a cuticlin protein. The *Dirofilaria immitis* nucleic acid molecule is useful as a probe to identify nucleic acid molecules, as a primer to produce nucleic acids, as a therapeutic reagent, e.g., gene therapy to inhibit cuticlin activity or production, or in a vaccine to prevent infection with helminth parasites. The cuticlin protein, antibodies raised against it, and inhibitory compounds of cuticlin may all be used in compositions to protect animals, especially mammals such as cats, dogs, and humans. The antibodies may be used to passively immunize an animal, or as reagents in assay to detect infection of helminths, or as tools to screen expression libraries to recover desired

CC proteins. They may also be used to target cytotoxic agents to the  
CC parasite and kill it directly. The present sequence represents the  
CC complement of DNA encoding a cuticlin polypeptide.

XX Sequence 1779 BP; 575 A; 302 C; 327 G; 575 T; 0 other;

Query Match 100.0%; Score 1161; DB 22; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATGATTCGCTTATTCCTTCTGTAACACTTATTCATTCGATTCGATTCG 60
    |||||||
DB 1613 ATGATGATTCGCTTATTCCTTCTGTAACACTTATTCATTCGATTCG 1554
    |||||||
QY 61 GTTGACAATGCTGCAAGGTGAGCCAGAAATGGAATGGACCACTTAATACATC 120
    |||||||
DB 1553 GTTGACAATGCTGCAAGGTGAGCCAGAAATGGAATGGACCACTTAATACATC 1494
    |||||||
QY 121 AATTTAATACAGTAAATGCAATGCAAGACATGTTATGGAAGCTTTATGATCAA 180
    |||||||
DB 1493 AATTTAATACAGTAAATGCAATGCAAGACATGTTATGGAAGCTTTATGATCAA 1434
    |||||||
QY 181 GAAGGTGCGGTAATGATGAAGGTGACGTCAGATTCGCGAAATTCACCTCCATTTGAT 240
    |||||||
DB 1433 GAAGGTGCGGTAATGATGAAGGTGACGTCAGATTCGCGAAATTCACCTCCATTTGAT 1374
    |||||||
QY 241 TCATGCAATGTCGCGTACACATCTCTGAATCCACGTCGATTTTGTGAACACACT 300
    |||||||
DB 1373 TCATGCAATGTCGCGTACACATCTCTGAATCCACGTCGATTTTGTGAACACACT 1314
    |||||||
QY 301 GTTGATATTCGTTTATTCATTCATTTGTACCAAGTTGATGTCGATTCGATTCGATCAA 360
    |||||||
DB 1313 GTTGATATTCGTTTATTCATTCATTTGTACCAAGTTGATGTCGATTCGATTCGATCAA 1254
    |||||||
QY 361 TCGTTTACATGAGACCTGATTAACAGATTAAGTACAGATGATGATTCGAATCACA 420
    |||||||
DB 1253 TCGTTTACATGAGACCTGATTAACAGATTAAGTACAGATGATGATTCGAATCACA 1194
    |||||||
QY 421 ACAGCTTTTCAACTCAAAATGTCGCGATCCGATATCCGTTATGAAATTTGGATGCT 480
    |||||||
DB 1193 ACAGCTTTTCAACTCAAAATGTCGCGATCCGATATCCGTTATGAAATTTGGATGCT 1134
    |||||||
QY 481 GGACCAACCGGTCACAGTTCATTTGCTATCATTTGGTCAGCAGTTTTCATTAATGG 540
    |||||||
DB 1133 GGACCAACCGGTCACAGTTCATTTGCTATCATTTGGTCAGCAGTTTTCATTAATGG 1074
    |||||||
QY 541 ACATGCGATTCGAAACCGTTGATCTTCTGCGGCTGTCATTCCTGCTTGTGAT 600
    |||||||
DB 1073 ACATGCGATTCGAAACCGTTGATCTTCTGCGGCTGTCATTCCTGCTTGTGAT 1014
    |||||||
QY 601 GATGTAAGCGGTATCTGGAATTTCTAATGCTGATGATGCTGCTGATTAATAT 660
    |||||||
DB 1013 GATGTAAGCGGTATCTGGAATTTCTAATGCTGATGATGCTGCTGATTAATAT 954
    |||||||
QY 661 TTGCTAAATTAATTTGGAATTCACAGATTTAATGGCTGGCAAGAGCTCCGATTAAC 720
    |||||||
DB 953 TTGCTAAATTAATTTGGAATTCACAGATTTAATGGCTGGCAAGAGCTCCGATTAAC 894
    |||||||
QY 721 AATATATCGGATGATCAGCTTTTCTATCAATGCCAGATCAGATTTACCATTAAGAA 780
    |||||||
DB 893 AATATATCGGATGATCAGCTTTTCTATCAATGCCAGATCAGATTTACCATTAAGAA 834
    |||||||
QY 781 CCAATATAGGAATGCTGTGACCACAATGTTCAAGAACACACAGATTCGAGCTGTTAAA 840
    |||||||
DB 833 CCAATATAGGAATGCTGTGACCACAATGTTCAAGAACACACAGATTCGAGCTGTTAAA 774
    |||||||
QY 841 ACAGGTGTCGCGCAGCAAAAACCTGCTGACGTCGCGCAATTCGTTTACCAAGAAAGA 900
    |||||||
DB 773 ACAGGTGTCGCGCAGCAAAAACCTGCTGACGTCGCGCAATTCGTTTACCAAGAAAGA 714
    |||||||
QY 901 TCTGCAAGACCGGAGATATCATTTGATGTAAGCAACTGATATCAACACCTTGAATTAAGC 960
    |||||||
DB 713 TCTGCAAGACCGGAGATATCATTTGATGTAAGCAACTGATATCAACACCTTGAATTAAGC 654

```

```

QY 961 GATGATTAATCAAGCTTTGCCAGTTGATTTAGTCACCGTCGACTTTCGCAACATTAATGA 1020
    |||||||
DB 653 GATGATTAATCAAGCTTTGCCAGTTGATTTAGTCACCGTCGACTTTCGCAACATTAATGA 594
    |||||||
QY 1021 CAACCTGTAATACCTGCTGCAAGTACAAATAATGATTCGATTCGATTCGATTCGATTCGA 1080
    |||||||
DB 593 CAACCTGTAATACCTGCTGCAAGTACAAATAATGATTCGATTCGATTCGATTCGATTCGA 534
    |||||||
QY 1081 ATGTTATGCGGTTTAACATTCGATTTGTCGCGCTCATTTATTAACATTCGTTTAAA 1140
    |||||||
DB 533 ATGTTATGCGGTTTAACATTCGATTTGTCGCGCTCATTTATTAACATTCGTTTAAA 474
    |||||||
QY 1141 TTTCGTCGAATCAGAGGCA 1161
    |||||||
DB 473 TTTCGTCGAATCAGAGGCA 453

```

## RESULT 5

ABL57372  
ID ABL57372 standard; cDNA; 1779 BP.

AC ABL57372;

DT 12-AUG-2002 (first entry)

DE Dirofilaria immitis cuticlin D1Cut-1A cDNA.

XX Cuticlin; helminth; parasite; D1Cut-1A; vaccine; gene therapy;

KW anthelmintic; gene; ds.

OS Dirofilaria immitis.

XX key Location/Qualifiers

FT CDS 167..1330

FT /tag= a

FT /product= "D1Cut-1A"

FT /note= "the CDS is also specifically claimed in

Claim 1"

XX US2002037294-A1.

XX 28-MAR-2002.

XX 20-MAR-2001; 2001US-0812642.

XX 01-JUN-1999; 99US-0323427.

XX (CHAN/) CHANDRASHEKAR R.

XX (MORA/) MORALES T H.

XX Chandrashekar R, Morales TH;

XX WPI: 2002-401306/43.

XX P-Psdb; ABB76278.

XX New parasitic helminth cuticlin proteins and nucleic acids isolated

XX form Dirofilaria immitis, useful for protecting animals from diseases

XX caused by parasitic helminths or for inhibiting molting of filarid

XX larvae in an animal -

XX Claim 1; Page 18-19; 32pp; English.

XX The present sequence is the coding strand of cDNA encoding novel

XX Dirofilaria immitis cuticlin D1Cut-1A protein (see ABB76278).

XX Nucleic acids comprising the complementary strand, the coding

XX sequence coding strand and its complementary strand are also

CC inhibit cuticlin activity, and methods of obtaining them.  
 CC Therapeutic compositions comprising such proteins, nucleic acids,  
 CC antibodies and/or inhibitors, including genetic vaccines,  
 CC recombinant virus vaccines and recombinant cell vaccines, are used  
 CC to protect animals from diseases caused by parasitic helminths, and  
 CC to inhibit the moulting of filarid larvae in an animal.

XX Sequence 1779 BP; 575 A; 327 C; 302 G; 575 T; 0 other;

Query Match 100.0%; Score 1161; DB 24; Length 1779;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATGATTCGCTTATTCCTTCTGCTACTACATTAATGCAATTCATTCGATTCGCG 60
DB 167 ATGATGATTCGCTTATTCCTTCTGCTACTACATTAATGCAATTCATTCGATTCGCG 226
QY 61 GTTGACAAATGGTGTGCAAGGTGAGCCAGAAATGGAATGTGGACCACTCAATACAAATC 120
DB 227 GTTGACAAATGGTGTGCAAGGTGAGCCAGAAATGGAATGTGGACCACTCAATACAAATC 286
QY 121 AATTTAATACAGTATGATTCGATTCGAAGACATGTTATGTGAAGTCTTATGATCA 180
DB 287 AATTTAATACAGTATGATTCGATTCGAAGACATGTTATGTGAAGTCTTATGATCA 346
QY 181 GAAGTTGCCGTATGATGAAGTGAGCGTCAAGTTGCCGAATTCACATTCATTTGAT 240
DB 347 GAAGTTGCCGTATGATGAAGTGAGCGTCAAGTTGCCGAATTCACATTCATTTGAT 406
QY 241 TCATGCAATGTGGCGGTACAGATCTCGAATCCAGCGGTATTTTAAACAACT 300
DB 407 TCATGCAATGTGGCGGTACAGATCTCGAATCCAGCGGTATTTTAAACAACT 466
QY 301 GTTGATATTCCTTATTCATTCATTAATTTGTACCAAGTTGATCGTCATATGAGTACA 360
DB 467 GTTGATATTCCTTATTCATTCATTAATTTGTACCAAGTTGATCGTCATATGAGTACA 526
QY 361 TCGTTTACATGGAAGCTATTAACAGTAAAGTAACTGACAGATTGAGGTATCGAAATCACA 420
DB 527 TCGTTTACATGGAAGCTATTAACAGTAAAGTAACTGACAGATTGAGGTATCGAAATCACA 586
QY 421 ACTGCTTTCAAACTCAATTTGCCGATGCCAGTATGCCGTTATGCAATTTTGGATGCT 480
DB 587 ACTGCTTTCAAACTCAATTTGCCGATGCCAGTATGCCGTTATGCAATTTTGGATGCT 646
QY 481 GGACCAACCGGTCAACAGTCAATTTGCTATCATTTGTCACAGTATATCAATAATGG 540
DB 647 GGACCAACCGGTCAACAGTCAATTTGCTATCATTTGTCACAGTATATCAATAATGG 706
QY 541 ACATGCGATTCGAAACCGTTGATCTTCTGCGCGTTGTCATTCCTGCTTTGTGAT 600
DB 707 ACATGCGATTCGAAACCGTTGATCTTCTGCGCGTTGTCATTCCTGCTTTGTGAT 766
QY 601 GATGTAACGCGTATCTGTGGAATTTCTAATGCTGATGATGCTTCTTATTAATAT 660
DB 767 GATGTAACGCGTATCTGTGGAATTTCTAATGCTGATGATGCTTCTTATTAATAT 826
QY 661 TTGCTAAATATTTTGAATATCCACAGATTTAATGGCTGGCCAGAAAGCTCAGATATC 720
DB 827 TTGCTAAATATTTTGAATATCCACAGATTTAATGGCTGGCCAGAAAGCTCAGATATC 886
QY 721 AAATATGCGGATCGATCAACAGCTTTCTATCAATGCCAGATCAGTATTAACATTAAGNA 780
DB 887 AAATATGCGGATCGATCAACAGCTTTCTATCAATGCCAGATCAGTATTAACATTAAGNA 946
QY 781 CCAATATAGCAATGTGTTGACACAAATGTTTGAACCAACAAGATTGGAGCTGTAA 840
DB 947 CCAATATAGCAATGTGTTGACACAAATGTTTGAACCAACAAGATTGGAGCTGTAA 1006
QY 841 ACAGGTGGTGGCGGAGCAAAACCTGCTGAGCTGGCAAACTTCGTTTACTGAAGAAAG 900
DB 1007 ACAGGTGGTGGCGGAGCAAAACCTGCTGAGCTGGCAAACTTCGTTTACTGAAGAAAG 1066

```

```

QY 901 TCTGAGAACCGGAGATATCATTTGATGACGAAGTATATCAACACCTTGAATTTAC 960
DB 1067 TCTGAGAACCGGAGATATCATTTGATGACGAAGTATATCAACACCTTGAATTTAC 1126
QY 961 GATGATATCAACGCTTTGGCATTTGATTTACGTCACCCGCACTTCTGCAATATGGA 1020
DB 1127 GATGATATCAACGCTTTGGCATTTGATTTACGTCACCCGCACTTCTGCAATATGGA 1186
QY 1021 CAACCTGATATCTGCTCAGTACAAATGGAATCTGATGTCACATTTGGCTTCTCA 1080
DB 1187 CAACCTGATATCTGCTCAGTACAAATGGAATCTGATGTCACATTTGGCTTCTCA 1246
QY 1081 ATGTTATGAGTTTAAAGCATTTGATTTGCTGCCGCTATTAATTAACATTTGTTAAA 1140
DB 1247 ATGTTATGAGTTTAAAGCATTTGATTTGCTGCCGCTATTAATTAACATTTGTTAAA 1306
QY 1141 TTTCTGCCAAATCAGAGGCA 1161
DB 1307 TTTCTGCCAAATCAGAGGCA 1327

```

#### RESULT 6

AAH24830  
 ID AAH24830 standard; DNA; 1372 BP.

XX AAH24830;

XX 22-AUG-2001 (first entry)

DE Nucleotide sequence of sense strand of cuticlin gene.

XX Cuticlin; gene therapy; vaccine; helminth parasite; ss.

XX Dirofilaria immitis.

OS key Location/Qualifiers

FT CDS 392..1207

FT /tag= a

FT /product= "cuticlin"

PN US6248329-B1.

XX 19-JUN-2001.

XX 01-JUN-1999; 99US-0323427.

XX 01-JUN-1998; 98US-0087435.

XX (CHAN/) CHANDRASHEKAR R.

XX (MORA/) MORALES T H.

XX Chandrashekar R, Morales TH;

XX WPI: 2001-396953/42.

XX P-PSDB: AAB84259.

XX Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin

XX protein is useful as a vaccine to prevent parasitic helminth infection

XX -

XX Example 1; Column 43-44; 29pp; English.

CC The specification describes a Dirofilaria immitis nucleic acid molecule,  
 CC that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid  
 CC molecule is useful as a probe to identify nucleic acid molecules, as a  
 CC primer to produce nucleic acids, as a therapeutic reagent, e.g., gene  
 CC therapy to inhibit cuticlin activity or production, or in a vaccine to  
 CC prevent infection with helminth parasites. The cuticlin protein,  
 CC antibodies raised against it, and inhibitory compounds of cuticlin may  
 CC all be used in compositions to protect animals, especially mammals such  
 CC as cats, dogs, and humans. The antibodies may be used to passively  
 CC immunize an animal, or as reagents in assay to detect infection of  
 CC helminths, or as tools to screen expression libraries to recover desired



CC proteins. They may also be used to target cytotoxic agents to the  
 CC parasite and kill it directly. The present sequence encodes a cuticlin  
 CC polypeptide.

XX Sequence 1372 BP; 456 A; 229 C; 281 G; 406 T; 0 other;

Query Match 43.4%; Score 504.2; DB 22; Length 1372;

Best Local Similarity 76.9%; Pred. No. 2.5e-131;

Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;

```

OY 1 ATGATGATTCGCTTATTCCTTCTGTAACACTTATG-CATTGCTTATTCGATTC 59
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 32 ATTATCATTTCTCTCTGTTATTCCTACTTTTCTACTTACCTAGCATCATGCC 91
OY 60 GGTTCACATGTTGTGACAGCTGACGACGAATGTAATGTGACCACTTCATTA 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 92 TATTGACAAATGTTGCGAAGTGAACCTGAATGTAATGTGCGGAGCTTCGAT 151
OY 120 CAATTTTAATACAGTAATGATTCGAAAGCAATGTTATGTAAGGCTTTATG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 152 CAATTTCAATACAGTAATGATTCGAAAGCAATGTTATGTAAGGACTTATG 211
OY 180 AGAAGTTCGCTTAATGATGAAAGTGAGCTCAAGTTCGGAATTCACCTTCAT 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 212 GATGATATGCTTCAGATGTAATGAGCGGAGTACGTGAATCGAATGGCAATG 271
OY 240 TTATGCAATGTTGCGGTACACGATCTGAAATCCAGTGTGATTTTGTAAACA 299
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 272 TTGCTGAATGTTGAAGCATCAGATCCTTAATCCGCGTGTGTTTGTAACTGT 331
OY 300 TGTTCATTCGTTGCTTATTCCTTATTCACAAATGATGATGATGAGTACA 359
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 332 AGTTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 391
OY 360 ATGCTTTTACATG-AAGTGATTAACACTTAGTGC-ACAGATGAGTATCGAATC 417
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 392 ATGCTTTTATGGAAGCTGATAGACGCTGATGCTGCTGTAAGTATGGAATG 451
OY 418 ACAACGCTTTTAACTCAATGTCGCCAGTCCGATGCGGTTATGAAATTTGGAT 477
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 452 ACTACAGCATTCCAAACTCAAGTGATCAATCCCGTATGATGATGATGATTTG 511
OY 478 GGTGACCAACCGTCAACAGTTCATTCATTCATTCATTCATTCATTCATTCAT 537
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 512 GGTGACCAACCGTCAACAGTTCATTCATTCATTCATTCATTCATTCATTCAT 571
OY 538 TGGACATGCGATTCGAAACCGTTCATTCATTCATTCATTCATTCATTCATTC 597
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 572 TGGACATGCGATTCGAAACCGTTCATTCATTCATTCATTCATTCATTCATTC 631
OY 598 GATGATGCTACGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 657
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 632 GATGATGCTACGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 691
OY 658 TATTGCTAAATTTGGAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 717
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 692 TATTGCTAAATTTGGAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 751
OY 718 TACAAATATGCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 777
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 752 TATTAATATGCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 811
OY 778 GAACCAATATGCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 837
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 812 GAACCAATATGCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 871
OY 838 AAAACAG 845
   ||||| |||||
Db 872 AATCTGG 879

```

RESULT 7  
 AAH24831/c

```

ID AAH24831 standard; DNA; 1372 BP.
XX
AC AAH24831;
XX
DT 22-AUG-2001 (first entry)
XX
DE Nucleotide sequence of antisense strand of cuticlin gene.
XX cuticlin; gene therapy; vaccine; helminth parasite; ss.
XX
OS Dirofilaria immitis.
OS
PN US6248329-B1.
PN
PD .19-JUN-2001.
PD
PF 01-JUN-1999; 99US-0323427.
PF
PR 01-JUN-1998; 98US-0087435.
PR
XX
PA (CHAN/) CHANDRASHEKAR R.
PA (MORA/) MORALES T H.
PI Chandrashekar R, Morales TH;
PI
XX WPI; 2001-396953/42.
XX
PT Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin
PT protein is useful as a vaccine to prevent parasitic helminth infection
PT
XX
PS Example 1; Column 43-46; 29pp; English.
PS
XX
CC The specification describes a Dirofilaria immitis nucleic acid molecule,
CC that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid
CC molecule is useful as a probe to identify nucleic acid molecules, as a
CC primer to produce nucleic acids, as a therapeutic reagent, e.g., gene
CC therapy to inhibit cuticlin activity or production, or in a vaccine to
CC prevent infection with helminth parasites. The cuticlin protein,
CC antibodies raised against it, and inhibitory compounds of cuticlin may
CC all be used in compositions to protect animals, especially mammals such
CC as cats, dogs, and humans. The antibodies may be used to passively
CC immunize an animal, or as reagents in assay to detect infection of
CC helminths, or as tools to screen expression libraries to recover desired
CC proteins. They may also be used to target cytotoxic agents to the
CC parasite and kill it directly. The present sequence represents the
CC complement of DNA encoding a cuticlin polypeptide.
XX
SO Sequence 1372 BP; 406 A; 281 C; 229 G; 456 T; 0 other;

Query Match 43.4%; Score 504.2; DB 22; Length 1372;
Best Local Similarity 76.9%; Pred. No. 2.5e-131;
Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;

OY 1 ATGATGATTCGCTTATTCCTTCTGTAACACTTATG-CATTGCTTATTCGATTC 59
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1341 ATTATCATTTCTCTCTGTTATTCCTACTTTTCTACTTACCTAGCATCATGCC 1282
OY 60 GGTTCACATGTTGTGACAGCTGACGACGAATGTAATGTGACCACTTCATTA 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1281 TATTGACAAATGTTGCGAAGTGAACCTGAATGTAATGTGCGGAGCTTCGAT 1222
OY 120 CAATTTTAATACAGTAATGATTCGAAAGCAATGTTATGTAAGGCTTTATGATCA 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1221 CAATTTCAATACAGTAATGATTCGAAAGCAATGTTATGTAAGGACTTATGATCA 1162
OY 180 AGAAGTTCGCTTAATGATGAAAGTGAGCTCAAGTTCGGAATTCACCTTCAT 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1161 GATGATATGCTTCAGATGTAATGAGCGGAGTACGTGAATGGCAATGGA 1102
OY 240 TTATGCAATGTTGCGGTACACGATCTGAAATCCAGTGTGATTTTGTAAACAAC 299
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1101 TTGCTGAATGTTGAAGCATCAGATCCTTAATCCGCGTGTGTTTGTAACTGT 1042

```



```

OY 300 TGTGTCATTTGCTTCATTCATTTATTTGTTACCAAGTTCATGTCGATTCAGTACA 359
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 1041 AGTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 982
OY 360 ATGCTTTTACATGAG-AAGCTGATTAACATGATGTC-ACAGATTCGATTCGAAATC 417
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 981 ATGTTTTTATATGGAAGCTGATTAAGCCCTTGTACTGCTGCTGATGATGTAATG 922
OY 418 ACAATGCTTTTCAAACTCAAAATGTCCTCCGATGCGCATGATCCGTTATGAAATTTGGAT 477
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 921 ACTACAGCATTCCTCAAACTCAAAATGTCATCCATGCGCTGATGATGATGATTTGGAA 862
OY 478 GGTGGACCAACCGGTACACCATTCATTTGCTATCATTTGCTGACCCAGTTATCATPAA 537
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 861 GGTGGACCAACCGGTGACCATTCATTTGCTATTCATTTGATGATGATGATGATGATG 802
OY 538 TGGACATGCGATTCGTAACCGTTGATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 597
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 801 TGGACATGCTGATTCAGAGACTACAGATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 742
OY 598 GATGATGTAACCGTATCTGTAATTCCTAAATGCTGATGATGATGCTGCTGCTGCTGCTG 657
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 741 GATGATGTAACCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 682
OY 658 TATTTGCTAAATATTTGGAATTCACACAGATTTAATGCTGCTGCTGCTGCTGCTGCTG 717
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 681 TATTTGCTAAATATTTGGAATTCATGATGATGATGATGATGATGATGATGATGATG 622
OY 718 TACAAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 621 TATTAATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 562
OY 778 GAACCAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 561 GAGCCATATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 502
OY 838 AAAACAGG 845
    ||| | | |
Db 501 AATATGCG 494

```

```

PA (MORA/) MORALES T H.
XX Chandrashekar R, Morales TH;
XX WPI: 2002-401306/43.
XX P-PSDB: ABB76279.
DR
PT New parasitic helminth cuticulin proteins and nucleic acids isolated
PT form Dirofilaria immitis, useful for protecting animals from diseases
PT caused by parasitic helminths or for inhibiting molting of filarid
PT larvae in an animal -
XX
XX Claim 1; Page 23; 32pp; English.
XX
XX The present sequence is the coding strand of cDNA encoding novel
XX Dirofilaria immitis cuticulin Dicut-1B protein (see ABB76279).
XX Nucleic acids comprising the complementary strand, the coding
XX sequence coding strand and its complementary strand are also
XX claimed. Dicut-1B cDNA was identified from a D. immitis larval
XX expressed sequence tag DNA sequence on the basis of homology to
XX Ascaris and Caenorhabditis elegans cut-1 cuticulin genes. The
XX encoded protein has 81% identity with the Ascaris homologue. The
XX invention provides D. immitis and Brugia malayi parasitic helminth
XX cuticulin proteins, nucleic acids, and methods of obtaining them.
XX Inhibit cuticulin activity, and methods of obtaining them.
XX Therapeutic compositions comprising such proteins, nucleic acids,
XX antibodies and/or inhibitors, including genetic vaccines,
XX recombinant virus vaccines and recombinant cell vaccines, are used
XX to protect animals from diseases caused by parasitic helminths, and
XX to inhibit the molting of filarid larvae in an animal.
XX
XX Sequence 1372 BP; 456 A; 229 C; 281 G; 406 T; 0 other;
XX
XX
XX Query Match 43.4%; Score 504.2; DB 24; Length 1372;
XX Best Local Similarity 76.9%; Pred No.2.5e-131;
XX Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;
OY 1 ATGATGATTCGCTTATTCCTTCTGCTACTACCTATTCG-CATTCCTTATTCGATTC 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 ATATCATATTCCTTCCTGTTATTCCTTATTCCTTATTCCTTATTCCTTATTCCTTATTC 91
OY 60 GGTGACATTCGTCGGAAGGTGACGCAATTTGATGATGATGATGATGATGATGATGATG 119
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 92 TATTCGATTCGTCGGAAGGTGACGCAATTTGATGATGATGATGATGATGATGATGATG 151
OY 120 CAATTTTATTCACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 179
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 152 CAATTTTATTCACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 211
OY 180 AGAAGTTCGCTTATTCGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 GATGATTCGCTTATTCGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 271
OY 240 TTTCATGCAATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 TTTCGTCATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 331
OY 300 TGTGTCATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 359
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 332 AGTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 391
OY 360 ATGCTTTTACATGAG-AAGCTGATTAACATGATGTC-ACAGATTCGATTCGAAATC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 ATGTTTTTATATGGAAGCTGATTAAGCCCTTGTACTGCTGCTGCTGCTGCTGCTGCTG 451
OY 418 ACAATGCTTTTCAAACTCAAAATGTCCTCCGATGCGCATGATCCGTTATGAAATTTGGAT 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 ACTACAGCATTCCTCAAACTCAAAATGTCATCCATGCGCTGATGATGATGATGATTTGGAA 511
OY 478 GGTGGACCAACCGGTGACCATTCATTTGCTATTCATTTGCTGACCCAGTTATCATPAA 537
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 512 GGTGGACCAACCGGTGACCATTCATTTGCTATTCATTTGCTGACCCAGTTATCATPAA 571

```

```
QY 538 TGGACATGGGATCTGGAACCGTTGATACCTTCGCGCGGTGNCATCTCCTGCTTGC 597
DB 572 TGGACATGTCATTCAGAGACTACAGATACATCTGTCATTAATGATCATGTGTTGTG 631
QY 598 GATGATGTAACGTCGATCTGGAATTCCTAAATGCTGATGATGCTGCTTGATAAA 657
DB 632 GATGATGTAACGTCGATCTGGAATTCCTAAATGCTGATGATGCTGCTTGATAAA 691
QY 658 TATTTGCTAAATTAATTTGGAATTCACAGAGATTTAAATGCTGCGCAAGAGCTCACGTA 717
DB 692 TATTTGCTAAATTAATTTGGAATTCACAGAGATTTAAATGCTGCGCAAGAGCTCACGTT 751
QY 718 TACAATATGCGGATGATGATACAGCTTTTCTATCAATGCGAGATCAGTATTAACATTA 777
DB 752 TATTAATATGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 811
QY 778 GAACCAATATGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
DB 812 GAGCAGATGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 871
QY 838 AAAACAG 845
DB 872 AAATCTGG 879

RESULT 9
AAH24839
ID AAH24839 standard; DNA: 892 BP.
AC AAH24839;
XX
XX 22-AUG-2001 (first entry)
DE Nucleotide sequence of a cuticlin polypeptide.
XX
XX Cuticlin: gene therapy; vaccine; helminth parasite; ss.
XX
XX Brugia malayi.
XX
XX
XX Key Location/Qualifiers
XX CDS 158..892
XX FT /*tag- a
XX FT /product- "cuticlin"
XX FT /note- "no termination codon given"
XX
XX US6248329-B1.
XX
XX 19-JUN-2001.
XX
XX 01-JUN-1999; 9905-0323427.
XX
XX 01-JUN-1998; 9805-0087435.
XX
XX (CHAN/) CHANDRASHEKAR R.
XX (MORA/) MORALES T H.
XX
XX Chandrashekar R, Morales TH;
XX
XX MPI; 2001-396953/42.
XX
XX P-PSDB; AAB84260.
XX
XX Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin
XX protein is useful as a vaccine to prevent parasitic helminth infection
XX
XX
XX Disclosure; Column 51-54; 29pp; English.
XX
XX The specification describes a Dirofilaria immitis nucleic acid molecule,
XX that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid
XX molecule is useful as a probe to identify nucleic acid molecules, as a
XX primer to produce nucleic acids, as a therapeutic reagent, e.g., gene
XX therapy to inhibit cuticlin activity or production, or in a vaccine to
XX prevent infection with helminth parasites. The cuticlin protein,
```

```
CC antibodies raised against it, and inhibitory compounds of cuticlin may
CC all be used in compositions to protect animals, especially mammals such
CC as cats, dogs, and humans. The antibodies may be used to passively
CC immunize an animal, or as reagents in assay to detect infection of
CC helminths, or as tools to screen expression libraries to recover desired
CC proteins. They may also be used to target cytotoxic agents to the
CC parasite and kill it directly. The present sequence encodes a cuticlin
CC polypeptide.
XX
XX SQ Sequence 892 BP; 290 A; 142 C; 168 G; 292 T; 0 other;
XX
XX Query Match 39.7%; Score 461; DB 22; Length 892;
XX Best Local Similarity 78.1%; Pred. No. 2.8e-119;
XX Matches 554; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
XX
QY 21 TTTCTGTACACACTTATGATGTCATTTGATGTCGCGGTGACATGTCGCAAG 80
DB 184 TTTGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 243
QY 81 TGAGCCGAATTTGAATGTGACCACTTCAATACATCAATTTTATACAGTAATGC 140
DB 244 TGAACCTGAATTAATGTGTCACACATCAATCTGTTAATTTAATCTCGAAATTC 303
QY 141 ATTGGAAGACATGTTTATGTAAGAGCTTTATGATCAAGAGCTTGCCTAATGATGA 200
DB 304 TTTTGAAGACATGTTTATGTAAGAGCTTTATGATCAAGAGCTTGCCTAATGATGA 363
QY 201 AGGTGAGACGTCAGTTCGCGCAATTTTCACTTCATTTGATTCATGATGTCGCGTAC 260
DB 364 AGGTGAGACGTCAGTTCGCGCAATTTTCACTTCATTTGATTCATGATGTCGCGTAC 423
QY 261 ACGATCTGTAATCCAGTGTGATTTTGTAAACAACAGTGTGATTCATTCATCC 320
DB 424 ACGTTCGTTAAATCCAGTGTGATTTTGTAAACAACAGTGTGATTCATTCATCC 483
QY 321 ATTATTTGTTACCAAGTGTGATTCGTCATATGAGTCAATGCTTTTACATGAGAGCTGA 380
DB 484 ACGTTCGTTAAATCCAGTGTGATTTTGTAAACAACAGTGTGATTCATTCATCC 543
QY 381 TAAACAGTTAGTGCACAGATTTGAGTATGTAATGTAATGTAATGTAATGTAATGTAAT 440
DB 544 TAAACAGTTAGTGCACAGATTTGAGTATGTAATGTAATGTAATGTAATGTAATGTAAT 603
QY 441 TGTCCCATGCAATATGTCGCTTATGAAATTTTGTGATGTCGCAACCGTCAACAGT 500
DB 604 GTTACCAATGCTGTGTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 663
QY 501 TCAATTTGCTATCATTTGTCAGCCAGTTTATCATTAATGAGATGCGATTCGTAACCGT 560
DB 664 CCAATATGCTATATTTGACACACCGGTTTATCATTAATGAGATGATGATGATGATGAT 723
QY 561 TGATACCTTCTGCGCGGTGTCATTCCTGCTTGTGATGATGATGATGATGATGATGAT 620
DB 724 TGATACCTTCTGCGCGGTGTCATTCCTGCTTGTGATGATGATGATGATGATGATGAT 783
QY 621 GGAATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
DB 784 TAAATTAATTAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
QY 681 TCCACAGATTTAATGCTGCGCAAGAGCTTATGATGATGATGATGATGATGATGATGAT 729
DB 844 TCCACAGATTTAATGCTGCGCAAGAGCTTATGATGATGATGATGATGATGATGATGAT 892

RESULT 10
AAH24840/c
ID AAH24840 standard; DNA: 892 BP.
AC AAH24840;
XX
XX 22-AUG-2001 (first entry)
XX
XX Nucleotide sequence of antisense strand of cuticlin open reading frame.
```



CC cuticlin nucleic acids. The invention provides D. immitis and  
 CC B. malayi parasitic helminth cuticlin proteins, nucleic acids,  
 CC antibodies, compounds that inhibit cuticlin activity, and methods  
 CC of obtaining them. Therapeutic compositions comprising such  
 CC proteins, nucleic acids, antibodies and/or inhibitors, including  
 CC genetic vaccines, recombinant virus vaccines and recombinant cell  
 CC vaccines, are used to protect animals from diseases caused by  
 CC parasitic helminths, and to inhibit the molting of filarial larvae  
 CC in an animal.

XX Sequence 892 BP; 290 A; 142 C; 168 G; 292 T; 0 other;

Query Match 39.7%; Score 461; DB 24; Length 892;

Best Local Similarity 78.1%; Pred. No. 2.8e-119;

Matches 554; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

```

OY 21 TTCTGTACACTTATTCGATTGTCTTATTCGATTCGGTTCAGACAGGTGCGAAG 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 TTGTCTATATATGATTAAGCAATATTAATGCTATTCATATGTAATGCTGCGAAG 243

OY 81 TGACCCGAATTTGAATGTGACCACTTCAATCAATCAATTTTAATACGTAATGC 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 TGAACCTGAATTTGAATGTGCTCAACATCACTGTTAATTTTACTCGAATCC 303

OY 141 ATTGGAAGACATGTTTATGGAAGGCTTTATGATGACAGAGGTTCCGTAATGATGA 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 TTTTGAAGACATGTTATATGTAAGATTTATACGTAATCAAGATTTCTGATGATGA 363

OY 201 AGGTGACGTCAAGTTGCCGAATTTCACTTCATTTGATTCATGCAATGTTGCCGTAC 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 364 AGGTGACGTCAAGTTGCCGAATTTCACTTCATTTGATTCATGTAATGTCGACGTAC 423

OY 261 ACGATCTCTGATGACAGTGTATTTTGTAAACAACATCTGTTCATTTGCTTCAATCC 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 424 ACGTTCGTTAATCCAGCTGAATATTTGTCAAGCTGTGTGTTAGCTTTCATCC 483

OY 321 ATTATTTGTTACCAAGTTGATCGTCATATCGATACATGCTTTTTCATGGAAGCTGA 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 484 ACATTTTATCCAAAAGTTGATGCAACATATCGATTCGAAATGCTTTTATCGAAGCTGA 543

OY 381 TAAACAGTTATGTCACACAGATTGAGTATCGAATCAACATCTTTTCAAACTCAAT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 544 TAAACAGTTATGTCACACAAATTTGAAGTTCCGAATGACACCTATTTGCTACAAAT 603

OY 441 TGTCCCATGCGCATGATGCCCTTATGAAATTTGGATGTCGACCAACCGTCAACCGT 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 604 GGTACCAATGCTGTGTGTATGATATGATTCGTGATGTCGACCAACCGTCAACCGT 663

OY 501 TCAATTTGCTATCATTTGTCAGCCAGTTATCATTAATGACATGCGATTTGGAACCGT 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 664 CCAATATGCTATATTTGACACACCGGTTTATCATTAATGACATGATTTCTGAAACGT 723

OY 561 TGATACCTTTCGCGCGGTTGTCATTCCTGCTTTGCGATGATGTAACGGTGAATGCT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 724 TGATACCTTTCGCGCTTTGTCATTCCTGCTTTGCGATGATGTAACGGTGAATGCT 783

OY 621 GGAATTTCTAATGCTGATGATGCTGCTTGATTAATTTTCTTAATTTTGAATA 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 784 TAAATTTATTAATGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 843

OY 681 TCCAAACAGATTATGCTGCGCAAGAACTCAGTATACAAATATGCG 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 844 TCCAACTGATTAATGCTGCGCAAGAACTCAGTATACAAATATGCG 892
  
```

RESULT 12  
 AAH24832  
 ID AAH24832 standard; DNA: 813 BP.

XX AAH24832;  
 AC  
 XX  
 DT 22-AUG-2001 (first entry)

```

DE Nucleotide sequence of a cuticlin polypeptide.
XX
XX Cuticlin; gene therapy; vaccine; helminth parasite; ss.
KW
XX
XX Dirofilaria immitis.
OS
XX
XX
XX Key
XX CDS
FT
FT 1.813
FT /*tag= a
FT /product= "cuticlin"
FT /note= "no termination codon given"
XX
XX
XX US6248329-B1.
XX
XX
XX 19-JUN-2001.
XX
XX 01-JUN-1999; 99US-0323427.
XX
XX 01-JUN-1998; 98US-0087435.
XX
XX (CHAN/) CHANDRASHEKAR R.
XX (MORA/) MORALES T H.
XX
XX Chandrashekar R, Morales TH;
XX
XX WPI: 2001-396953/42.
XX P-PSDB: AAB84259.
XX
XX
XX Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin
XX protein is useful as a vaccine to prevent parasitic helminth infection
XX
XX
XX Example 1: Column 45-48; 29pp; English.
XX
XX
XX The specification describes a Dirofilaria immitis nucleic acid molecule,
XX CC that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid
XX CC molecule is useful as a probe to identify nucleic acid molecules, as a
XX CC primer to produce nucleic acids, as a therapeutic reagent, e.g., gene
XX CC therapy to inhibit cuticlin activity or production, or in a vaccine to
XX CC prevent infection with helminth parasites. The cuticlin protein,
XX CC antibodies raised against it, and inhibitory compounds of cuticlin may
XX CC all be used in compositions to protect animals, especially mammals such
XX CC as cats, dogs, and humans. The antibodies may be used to passively
XX CC immunize an animal, or as reagents in assay to detect infection of
XX CC helminths, or as tools to screen expression libraries to recover desired
XX CC proteins. They may also be used to target cytotoxic agents to the
XX CC parasite and kill it directly. The present sequence encodes a cuticlin
XX CC polypeptide.
XX
XX
XX Sequence 813 BP; 247 A; 148 C; 164 G; 254 T; 0 other;
XX
XX
XX Query Match 26.7%; Score 310.4; DB 22; Length 813;
XX Best Local Similarity 70.7%; Pred. No. 5.9e-77;
XX Matches 455; Conservative 0; Mismatches 161; Indels 28; Gaps 2;
XX
XX
XX 372 GGAAGCTGATTAACAGTTAGTC-ACAGATTGAGGTCTCGAATTCACACTGCTTTTC 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14 GAAAGCTGATTAAGACCGTTTACTGCTCTTGAAGTCTGGAATGACTACACCATCC 73

OY 431 AAACGTCAATTTGCCCGGTCAGTATGCGCTTTGAAATTTGGATGTCGACCAACCG 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 74 AAACGTCAATTTGCCCGGTCAGTATGCGCTTTGAAATTTGGATGTCGACCAACCG 133

OY 491 GTCAACAGTTCAATTTGCTATCATTTGTCAGCCAGTTTATCATTAATGACATGCGATT 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 GTGACACCTGTGATTTGCAATGATTTGGAATGATGATATCAACAATGACATGATTT 193

OY 551 CTGAACCGTGTATCTTTCGCGGTTGTCATTCCTGCTTTGTCATGATGATGATGACG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 CAGAGACTACAGATACATCTGTCATTTAGTATTCATTCATGCTGTGATGATGAAAG 253

OY 611 GTGATACGTGGAATTTTAATGCTGATGATGATGCTCTGATTAATTTGTAATA 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Db 254 GTGATGCTGAGTTCGATGGAAGAAGATGCTTTGGACAAATATTACTCAARA 313  
 QY 671 ATTGGAATATCCACAGATTTATGCTGGCCAGAGCTCAGTATCAATATGCGG 730  
 Db 314 ATTGGAATATATATACAGATTTATGCTGGCCAGAGCTCAGTATCAATATGCGG 373  
 QY 731 ATGATGACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAGAACAATAGCG 790  
 Db 374 ATGATGACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAGAACAATAGCG 433  
 QY 791 AATGTTGGTTCACACATTTTCCAGACCAAGGATTCG----- 830  
 Db 434 AATGTTGGTTCACACATTTTCCAGACCAAGGATTTGCTCCATTAATTCGACAG 493  
 QY 831 -----ACCTGTTAAACAGGTGTCGCCAGCAAAACCTGTCGACGTCGCAACTTC 883  
 Db 494 GATTGCTGCTGTAAATATGCTGCGTGCACAGCTCCAGAGTTCCTGCTTCTCCAC 553  
 QY 884 GTTACTCTCAAGAAAGATCTGCAGAACCGAGATATCATTTGATGACAGACTGATATCA 943  
 Db 554 GATTGATCAAGAGCGATCAATTAATTCGATTAATGCGTGGACGTCAGTACGGTTTAA 613  
 QY 944 ACACCCCTTGAATTTAGCGATGAATCAAGCTTTGCGAGTTGAT 987  
 Db 614 GCACGGTTGATATTAACCGAGAGAAATCCGAACTTCTCAGCAAT 657

RESULT 13  
 AAH24833/C  
 ID AAH24833 standard; DNA; 813 BP.

AAH24833;

22-AUG-2001 (first entry)

Nucleotide sequence of antisense strand of cuticlin open reading frame.

Cuticlin; gene therapy; vaccine; helminth parasite; ss.

Dirofilaria immitis.

US6248329-B1.

19-JUN-2001.

01-JUN-1999; 99US-0323427.

01-JUN-1998; 98US-0087435.

(CHAN/) CHANDRASEKHAR R.

(MORA/) MORALES T H.

Chandrasekar R, Morales TH;

WPI: 2001-396953/42.

Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin protein is useful as a vaccine to prevent parasitic helminth infection

Example 1; Column 49-50; 29pp; English.

The specification describes a Dirofilaria immitis nucleic acid molecule, that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid molecule is useful as a probe to identify nucleic acid molecules, as a primer to produce nucleic acids, as a therapeutic reagent, e.g., gene therapy to inhibit cuticlin activity or production, or in a vaccine to prevent infection with helminth parasites. The cuticlin protein, antibodies raised against it, and inhibitory compounds of cuticlin may all be used in compositions to protect animals, especially mammals such as cats, dogs, and humans. The antibodies may be used to passively immunize an animal, or as reagents in assay to detect infection of helminths, or as tools to screen expression libraries to recover desired

CC proteins. They may also be used to target cytotoxic agents to the parasite and kill it directly. The present sequence is the complement of an open reading frame encoding a cuticlin polypeptide.

Sequence 813 BP; 254 A; 164 C; 148 G; 247 T; 0 other;

Query Match 26.7%; Score 310.4; DB 22; Length 813;

Best Local Similarity 70.7%; Pred. No. 5.9e-77;

Matches 455; Conservative 0; Mismatches 161; Indels 28; Gaps 2;

QY 372 GGAAGCTGATTAACAGTTAGTGC-ACAGATGAGTATCTGAATACAACTGCTTTG 430  
 Db 800 GAAAGCTGATTAACAGTTAGTGC-ACAGATGAGTATCTGAATACAACTGCTTTG 741  
 QY 431 AACTCAATTTGCTCCGATGAGTATGCGCTTATGAATTTGGATGGACCAACCG 490  
 Db 740 AACTCAATTTGCTCCGATGAGTATGCGCTTATGAATTTGGATGGACCAACCG 681  
 QY 491 GTCAACAGTTCAATTTGCTATCATTTGCTACAGCCAGTTATCATTAATGACATGAT 550  
 Db 680 GTCAACAGTTCAATTTGCTATCATTTGCTACAGCCAGTTATCATTAATGACATGAT 621  
 QY 551 CTGAACCGTTGATATCTTCTGCGGCTGTCATTCCTGCTTTGCTGATGATGATG 610  
 Db 620 CAGAGCTACAGATCATTTCTGCTATGATGATGATGATGATGATGATGATGATG 561  
 QY 611 GTGATACTGTTGAATTTCTAATGCTGATGATGATGATGATGATGATGATGATG 670  
 Db 560 GTGATGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501  
 QY 671 ATTTGGAATATCAACAGATTTAATGCTGCGCAAGAAAGCTCAGTATCAATATGCGG 730  
 Db 500 ATTTGGAATATTAATACAGATTTAATGCTGCGCAAGAAAGCTCAGTATTAATATGACG 441  
 QY 731 ATGATACAGCTTTTCTATCAATGCGAGATGATGATGATGATGATGATGATGATG 790  
 Db 440 ATGATACAGCTTTTCTATCAATGCGAGATGATGATGATGATGATGATGATGATG 381  
 QY 791 AATGTTGCTGACACAAATGTTGACAGCAACAGAGATGCTG----- 830  
 Db 380 AATGTTGCTGACACAAATGTTGACAGCAACAGAGATGCTG----- 830  
 QY 831 -----AGCTGTTAAACAGGTGTCGCCAGCAAAACCTGTCGACGTCGCAACTTC 883  
 Db 320 GATTGCTGCTGTAAATATGCTGCTGCAACGCTCCAGAGCTTCTGCTTCTCCAC 261  
 QY 884 GTTACTCAAGAAAGATCTGCAGAACCGGAGATATCATTTGATGATGATGATGATCA 943  
 Db 260 GATTGATCAAGAGCGATCAATTAATTTGATTAATGCGTGGACGTCAGTACCGGTTTAA 201  
 QY 944 ACACCCCTTGAATTTAGCGATGAATCAAGCTTTGCGAGTTGAT 987  
 Db 200 GCACGGTTGATATTAACCGAGAGAAATCCGAACTTCTCAGCAAT 157

RESULT 14  
 AAV21209/C  
 ID AAV21209 standard; DNA; 1664976 BP.

AAV21209;

10-NOV-1998 (first entry)

Methanococcus jannaschii circular chromosome.

Methanococcus jannaschii; methanogenic archaeon; circular chromosome; genome; autotrophic; extrachromosomal element; identification; ds.

Methanococcus jannaschii.

W09807830-A2.

26-FEB-1998.

```
XX 22-AUG-1997; 97MO-US14900.
XX PR 22-AUG-1996; 96US-0024428.
XX PA (GENO-) INST GENOMIC RES.
XX PA (UNIV ) UNIV ILLINOIS FOUND.
XX PA (UJJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX PI Bult CJ, Smith HO, Venter JC, White OR, Moese CR.
XX DR WPI; 1998-169145/15.
XX PT Complete genome sequence of methano-genic archaeon, Methanococcus
XX PT jannaschii - useful in identification of M. jannaschii genome
XX PT fragment
XX PS Claim 13; Page 152-585; 614pp: English.
XX CC The present sequence represents the complete 1.66-megabase pair genome
XX CC sequence of the Methanococcus jannaschii circular chromosome. The
XX CC present invention describes M. jannaschii open reading frames from the
XX CC genome sequence. The invention also describes a computer based system
XX CC for identifying fragments of the M. jannaschii genome that are
XX CC homologous to target nucleotide sequences, comprising: (a) data storage
XX CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550
XX CC bp sequence (see AAV21209, AAV21210 and AAV21211), or a nucleotide
XX CC sequence at least 99.98 identical to it; (b) search means for comparing a
XX CC target sequence to the nucleotide sequence of the data storage means to
XX CC identify a homologous sequence, and (c) retrieval means for obtaining
XX CC the homologous sequence. The method, which is based on whole genome
XX CC random sequencing of an autotrophic archaeon M. jannaschii, the genome
XX CC of which consists of 3 physically distinct elements, a large circular
XX CC chromosome (the 1664976 bp sequence given in AAV21209), a large circular
XX CC extra-chromosomal element (the 58407 bp sequence given in AAV21210), and
XX CC a small circular extra-chromosomal element (the 16550 bp sequence given
XX CC in AAV21211), can be used in the identification of M. jannaschii genome
XX CC fragment.
XX SQ Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other;
XX
Query Match 3.4%; Score 40; DB 19; Length 1664976;
Best Local Similarity 47.9%; Pred. No. 20;
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 667 AATTAATTTGGAATATCCACAGATTTAATGGCTGGCCAGAGCTCACGTAATACAAATAT 726
DB 684806 ATTAATTTGGAATATTTTAATCTCTTAATTAATGTTCACTTCTCAATCCTTCCTCA 684747
QY 727 GCGGATGATCAGACGCTTTCTATCAATGCCAGATCCAGTATACCAATTAAGAACAAT 786
DB 684746 CTGCAATTTGAGCTGTTGGTGTGATTTGACAGCAATCCGAGTTTCTCTCAACTCAACATTC 684687
QY 787 AGCGAATGTGTTCGACCAATATTTTCAGAACACACAGAGATTCGAGCTGTTAAAAAGGT 846
DB 684686 TGGTATTTTCGAGAGCAAGATGTCGCAATCTTAAACATTAAGAGCTGTTTCTATGCT 684627
QY 847 GGTGCGCCAGCAAACTGCTGCGAGCTGCCACTTGTCTTACTCAAGAAAAAGATCTGCA 906
DB 684626 GCTAATGCCATATTTCTTTTGTAGCGCGCAATATTTCTTAAAGTTTATGATTTCTGTA 684567
RESULT 15
ABL32426/C
ID ABL32426 standard; DNA; 9810 BP.
XX
XX ABL32426;
XX AC
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 399.
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;
```

```
KW antiarteriosclerotic; antianaemic; cyostatic; neutropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX OS Homo sapiens.
XX PN WO20020928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP07537.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIC-) EPIGENOMICS AG.
XX PT Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful
XX PT for diagnosis and treatment of diseases associated with abnormal
XX PT cytosine methylation -
XX PS Claim 1; SEQ ID NO 399; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention.
XX SQ Sequence 9810 BP; 2770 A; 79 C; 1978 G; 4983 T; 0 other;
XX
Query Match 3.4%; Score 39.4; DB 24; Length 9810;
Best Local Similarity 50.3%; Pred. No. 2.8;
Matches 97; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 651 TGATTAATATTTTCTTAATATTTGCAATATCCAGATTTAATGGCTGGCCAGAAGC 710
DB 5409 TATTCGAATCTCTCAATATATATACGAAATTTCAAAAAATTAATTCGACTCCAAACATA 5350
QY 711 TCACGTATCAATATATGGGATCGATCAGACGCTTTCTATCAATGCCAGATGATATTAC 770
DB 5349 ATTAATAAAAAAACAATACGTATCTTAATCTTAAACCTTTAAACCAAAACAATTTCTATTTAA 5290
QY 771 CATTAAAGAACCAATATGCGAATGTGTTCGACCAATGTTCAGACCAACAAGATTGCG 830
DB 5289 TTAATTAATAAACCACTAATAATCTTTATTCAAAAACATCAATCAACTACATCAAAAAAA 5230
QY 831 AGCTGTTAAACA 843
DB 5229 AACTAACAAACA 5217
RESULT 16
AAAX9556/C
ID AAAX9556 standard; DNA; 14066 BP.
XX
XX AAAX9556;
XX AC
XX DT 05-OCT-1999 (first entry)
XX DE Nucleic acid sequence from U. urealyticum.
```

KM Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;  
 KM human urogenital tract; pregnancy; neonatal disease; drug therapy;  
 KM suppurative arthritis; ss.  
 XX  
 OS Ureaplasma urealyticum.  
 XX  
 PN W09939007-A1.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 29-JAN-1999; 99WO-US01972.  
 XX  
 PR 30-JAN-1998; 98US-0073189.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Cassell GH, Chen EX, Glass JT, Glass JS, Helner CR;  
 PI Letkowitz E;  
 XX  
 DR WPI: 1999-469343/39.  
 XX  
 PT Detection of Ureaplasma urealyticum using novel genes, probes and  
 PT primers  
 XX  
 PS Claim 1; Page 41-46; 110pp; English.  
 XX  
 CC The present invention provides methods for the detection and diagnosis  
 CC of Ureaplasma urealyticum infection. It provides novel genes  
 CC (AA99501-681) that can be used as a source of primers and probes for the  
 CC detection and/or quantification of U. urealyticum in a biological  
 CC sample. The probes that can be used in the method of the invention by  
 CC forming target:probe complex is complementary to a region selected from  
 CC one of the 181 nucleic acid sequences (AA99501-681). U. urealyticum is  
 CC an opportunistic pathogen of the human urogenital tract that is a  
 CC significant cause of adverse pregnancy outcome, neonatal disease, and  
 CC suppurative arthritis. As the infections are commonly asymptomatic, it is  
 CC important to have specific and sensitive methods for detecting their  
 CC presence in a patient. Also, as the pathogen has no current antibiotic  
 CC directed specifically against it, it would be advantageous to isolate and  
 CC detect gene sequences which are unique to it, and utilise these as a  
 CC basis for diagnosis of U. urealyticum infection as well as to develop new  
 CC and improved drug therapies. The present invention provides such novel  
 CC polynucleotide sequences (AA99501-681).  
 XX  
 SQ Sequence 14066 BP; 4508 A; 1702 C; 1539 G; 6317 T; 0 other;  
 QY  
 Query Match 3.3%; Score 38.8; DB 20; Length 14066;  
 Best Local Similarity 57.4%; Pred. No. 4.8;  
 Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
 QY 81 TGAGCCAGAAATGTAATGTGCACTTCATATACATCAATTTTAATACAGTATATGC 140  
 DB 9196 TAAACCAAAATGTAATGTGCACTTCATATACATCAATTTTAATACAGTATATGC 9137  
 QY 141 ATTGCAAGCATGTTTATGTGTAAGGCTTTATGATCAAGAAGTTCGCGTATGATGA 200  
 DB 9136 TATTAAACAGTATGACATATATGCTCAAGTACATTTTAAGCTTAATAAACAATATGA 9077  
 QY 201 AG 202  
 DB 9076 AG 9075  
 RESULT 17  
 ABK39779/C  
 ID ABK39779 standard; CDNA; 6799 BP.  
 XX  
 AC ABK39779;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE cDNA encoding full length L1428P.  
 XX

KM Lung tumour; cancer; T cell; immune response stimulator;  
 KM cytosolic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200204514-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 10-JUL-2001; 2001WO-US22058.  
 XX  
 PR 11-JUL-2000; 2000US-0614124.  
 PR 29-AUG-2000; 2000US-0651563.  
 PR 08-SEP-2000; 2000US-0658824.  
 PR 26-SEP-2000; 2000US-0671325.  
 PR 06-OCT-2000; 2000US-0677419.  
 PR 30-OCT-2000; 2000US-0702705.  
 PR 13-DEC-2000; 2000US-0736457.  
 PR 03-MAY-2001; 2001US-0849626.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Matanabe Y, Henderson RA, Johnson JC, Retter KM;  
 PI Marnerakis M, Carter D, Fanger GR, Vedrick TS, Bangur CS;  
 PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
 XX  
 DR WPI: 2002-164634/21.  
 DR P-PSB; AA085654, AA085655, AA085656, AA085657, AA085658, AA085659,  
 DR AA085660.  
 XX  
 PT Novel polynucleotide encoding a lung tumour polypeptide useful for  
 PT stimulating and/or expanding T cells specific for a tumour protein  
 XX  
 PS Claim 1; SEQ ID NO 1883; 223pp; English.  
 XX  
 CC The invention describes an isolated polynucleotide and polypeptide  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein for determining the presence of a cancer in a patient. A  
 CC composition containing the polynucleotide and/or polypeptide is useful  
 CC for treating a lung cancer in a patient. The polypeptide is useful for  
 CC removing tumour cells from a biological sample. The polynucleotide is  
 CC also useful as probe or primer to detect the level of mRNA encoding a  
 CC tumour protein. This sequence encodes a lung tumour associated protein  
 CC or protein fragment, described in the method of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 6799 BP; 1801 A; 1865 C; 1539 G; 1594 T; 0 other;  
 QY  
 Query Match 3.3%; Score 38; DB 24; Length 6799;  
 Best Local Similarity 59.1%; Pred. No. 5.8;  
 Matches 65; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
 QY 235 TTGTGATCATGCAAGTGGCGGTACAGATCTGGAATCAACGATATTTTGTACAA 294  
 DB 5002 TTTGTACCAAGCAAAATGAGTCCAGCGAGCTCAATCGGCATGCTTCTGTGCA 4943  
 QY 295 ACAAGCTGTGATCTTTCGTTTCATCATTTATTTACCAAGTTGATCG 344  
 DB 4942 GTACCTGTGCTGATCTTATTTCTAGATGATTTATCTTTATTTATGG 4893  
 RESULT 18  
 ABQ70961/C  
 ID ABQ70961 standard; DNA; 6242 BP.  
 XX  
 AC ABQ70961;  
 XX  
 DT 29-AUG-2002 (first entry)  
 XX  
 DE Listeria monocytogenes 4b contig DNA sequence #903.  
 DE



XX Antibacterial; Listeria; food contamination; mutational analysis;  
KM infection; ds.  
XX  
OS Listeria monocytogenes 4b.  
XX  
PN WO200228891-A2.  
XX  
PD 11-APR-2002.  
PF 04-OCT-2001; 2001WO-FR03061.  
PR 04-OCT-2000; 2000FR-0012697.  
XX  
PA (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
PI Kunst F, Glaser P;  
PI  
DR WPI; 2002-332479/37.  
XX  
PT New genomic sequences from Listeria species, useful for detection,  
PT treatment and prevention of infection, also related polypeptides,  
PT antibodies and modulators -  
XX  
PS Claim 14; SEQ ID 3774; 180pp; French.  
XX  
CC The present invention relates to nucleic acid sequences  
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes  
CC and primers for identification and/or detection of Listeria (e.g. as  
CC contaminants in foods, or mutational analysis) and for analysis of  
CC gene expression. Proteins encoded by the nucleic acid sequences can be  
CC used to screen for compounds that modulate gene expression, replication  
CC and pathogenicity of Listeria (potential therapeutic agents), also for  
CC treating infections by Listeria, and are useful as immunogens in  
CC anti-Listeria vaccines.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 6242 BP; 1954 A; 1060 C; 1164 G; 2064 T; 0 other;  
Query Match 3.2%; Score 37.2; DB 24; Length 6242;  
Best Local Similarity 53.4%; Pred. No. 9.3;  
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
QY 821 AAGGATTCGAGCTGTTAAAAACAGGTGGTCCCGACCAAACTGCTGCAGCTGGCGAAC 880  
DB 3172 AAGAAAGTGGAAAGCGTTACAGCGCAATGCTTCGGTATCAATGATGAGCTTCCGCA 3113  
QY 881 TTGGTTACTCAGAAAGATCTGCAGAACCGGAGATATCATTTGATGTCAGAACTGATA 940  
DB 3112 TTATTTTAAAGTCGAAAGAAAGAGCGAGTCGGGAAAAACATTCCTATATCGCAACATCA 3053  
QY 941 TCAACACCCCTTGAATTAAGCATGAT 966  
DB 3052 AAGTAACCTCTGAAGTTGTGTGAT 3027  
RESULT 19  
AA216979/c  
ID AA216979 standard; cDNA; 783 BP.  
XX  
AC AA216979;  
XX  
DT 12-OCT-1999 (first entry)  
XX  
DE Human gene expression product cDNA sequence SEQ ID NO:4449.  
XX  
KM Human; gene: gene expression product; diagnosis; therapy; probe;  
KM detection; mapping; tissue typing; profiling; forensic; cancer;  
KM genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX

OS Homo sapiens.  
XX  
PN WO9938972-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 28-JAN-1999; 99WO-US01619.  
PR 03-APR-1998; 98US-0080666.  
PR 28-JAN-1998; 98US-0072910.  
PR 24-FEB-1998; 98US-0075954.  
PR 31-MAR-1998; 98US-0080114.  
PR 03-APR-1998; 98US-0080515.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
PI Escobedo J, Garcia PD, Garcia V, Gliese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lamson G, Lesnikowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX  
DR WPI; 1999-494092/41.  
XX  
PT Novel human genes and their expression products which are  
PT differentially expressed in different cell types  
XX  
PS Claim 1; Page 2107; 2479pp; English.  
XX  
CC The present invention describes a library of human polynucleotides  
CC comprising the sequences given in AA212532 to AA217779. Also described is  
CC a method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein), and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX  
SQ Sequence 783 BP; 243 A; 176 C; 145 G; 197 T; 22 other;  
Query Match 3.2%; Score 37; DB 20; Length 783;  
Best Local Similarity 53.0%; Pred. No. 4.1;  
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 588 CTGCTTTGCGATGATGTAACGGTATGTAAGTCTTAATCTGATGATGATGTC 647  
DB 392 CTGCTGAGAGAGAGAACTGATGAGTATGCAATTAAGTCAATTCCTGTTAACTGGC 333  
QY 648 TCTTGATTAATTTTGTAAATATTTGGAAATATCCACAGATTTATGCTGGCCAGA 707  
DB 332 AGTTTCAATATGTCATTAATTAACCATTCGATGATGATGAGATGAGCAAG 273  
QY 708 AGCTACGATACAAATATCGGATGAT 736  
DB 272 GGGTCCGTTTATCATCTTCTTGTAT 244  
RESULT 20  
AA299054/c







	Query Match	Best Local Similarity	Score 37;	DB 24;	Length 1680;
	Matches 82;	Conservative	0;	Mismatches 75;	Indels 0;
					caps
Oy	69	TGCTTCGAAGTGGACGAGAAATTTGATTTGGACCAACTTCATTAACAATCAATTTTAA	128		
Db	1030	TGCTTTTCGAAGAAAAGACGCGGATGACTTTTTCAGAACTTTTATGAGAGCGGATATTGA	971		
Oy	129	TACACGTAAATGATTCAGAGACATGTTTATGTGAAGGCTTTATGATCAAGAAAGTTG	189		

CC	corresponding gene of <i>A. thaliana</i> ecotype Columbia (see AA089385).
CC	Gene silencing is useful as a molecular tool for regulating gene expression.
CC	
XX	
50	Sequence 6571 BP; 2060 A; 1347 C; 1470 G; 1694 T; 0 other;
	Query Match 3.2%; Score 36.8; DB 22; Length 6571;
	Best Local Similarity 48.1%; Pred. No. 12;
	Matches 104; Conservative 0; Mismatches 112; Indels 0; Gaps

QY	743	TTTTCTATCAATGCCAGATCAGTATTTACCATTTAAAGAACCCAAATACGAGATGTGTTGCAC	802
Db	6564	TTTTTTTTTTTTAAAGTTAAATACGATAGTATTTTCATCATCATCAATTTGGTGTGTGATACAGA	6505
QY	803	CACAAATGTCGAGAACCAACAGAGATTTGGAGCCTTTAAACAGSTGCTGCCGACGACAAAC	862
Db	6504	CTGATTTTATATATACATGAAAGTGTGAACCGTTGCGCTTTCCAAAGAGAGACCTTAGGCAACG	6445
QY	863	CTGCTGCAGCTGCGCCACACTCTTTTACCTCAAGAAAGATCTCAGAACCGAGATATATCA	922
Db	6444	AGTCGGTTCCTCGACTCAGCTCAAAATTAGATTAAACGACAGAAATTAGTGAACCTTACTATACA	6385
QY	923	TTGAGTGTACGAACTGATATCAACACCCCTTGAATTA	958
Db	6384	TATATATATATACTATATATTAACCTCTCTAAATATCA	6349

XX	RESULT 27
XX	AAI65681
XX	ID AAI65681 standard; DNA; 1527 BP.
XX	AAI65681;
XX	03-JAN-2002 (first entry)
XX	Nucleotide sequence comprising a phage resistance mechanism.
XX	Phage resistance mechanism; lactic acid bacteria; cheese; yogurt;
XX	butter; sauerkraut; ss.
XX	Streptococcus thermophilus.
XX	OS
XX	Key
XX	Location/Qualifiers
XX	CDS 272..1456
XX	/*tag= a
XX	MO200170990-AL.
XX	27-SEP-2001.
XX	21-MAR-2001; 2001WO-FR00836.
XX	21-MAR-2000; 2000FR-0003578.
XX	(SUDD ) SKW NATURE PROD HOLDING FRANCE SAS.
XX	Prevots F, Tolou S;
XX	WPI; 2001-611505/70.
XX	DR P-PSDB; AAG79122.
XX	DR
XX	New nucleic acid for imparting phage resistance to lactic acid bacteria
XX	used in food preparation comprises a polynucleotide that encodes a
XX	phage-resistance mechanism -
XX	Claim 1; Page 16-17; 30pp; French.
XX	The present nucleic acid sequence comprises a phage resistance
XX	mechanism. Plasmids that contain the phage resistance mechanism
XX	nucleic acids are used to impart phage resistance to industrially
XX	important strains of lactic acid bacteria. These recombinant strains
XX	are then used for preparation of cheese, yogurt, butter, sauces and
XX	sauerkraut.
XX	Sequence 1527 BP; 612 A; 148 C; 259 G; 508 T; 0 other;

	Query Match	3.2%	Score 36.6;	DB 22;	Length 1527;	
	Best Local Similarity	47.2%;	Pred. No. 7.2;			
	Matches 111; Conservative	0;	Mismatches 124;	Indels 0;	Gaps 0;	
OY	592 TTTGCGATGATGCAGCCTCATCTGTGGAAATTCTTAAGTCGATGGANGTCTCTT	651				

Db	533	TTTGAGAGCAATTCGTCCTCAGTTCTTTTGAAAACAATTAACAGTTGAGACATGAAATT	5922
Qy	652	GATAAATATTTTCTAAATATTTTGGAAATATCCAAGATTAAATGCGTGCCCAAGAAGCT	7111
Db	593	GATGATTAAGTATTAAGAAATGACATTCACACTCCATATTTTAACTGAAAGTTTACT	6522
Qy	712	CACGATATCAAAATATGCGGATCGATCAACAGCTTTTATCAATGCCAGTCAGTATTACC	7711
Db	653	TCAAAAAAGAAAGATATTTCAAAACCCTAAGTTGTTTGAACAAATATCTACTAGG	7122
Qy	772	ATTAAAGACCAATTAGCGATGTGTTGACACAAATGTTCAAGACCACAAGAT	826
Db	713	ATTATGATGATTAATACCAATTAATCGATCTGCGAGATGTTATTAATAGAT	767

RESUL	28
ABL32500	
ID	ABL32500 standard; DNA; 8342 BP.
XX	
AC	ABL32500;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human immune system associated gene SEQ ID NO: 473.
XX	
KW	Human immune system disease; cytosine methylation; antiasthmatic;
KW	antiartherosclerotic; antiamebic; cytosolic; nootropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW	antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW	gene; ds.
OS	Homo sapiens.
XX	
PN	WO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIC-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
XX	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation
XX	
PS	Claim 1; SEQ ID NO 473; 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	mucular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
XX	
Sequence	8342 BP; 2037 A; 159 C; 2425 G; 3721 T; 0 other;

Query Match 3.2%; Score 36.6; DB 24; Length 8342;  
 Best Local Similarity 50.9%; Pred. NO. 16;  
 Matches 87; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Db 1782 TTTTGGTAATAGTAATTTTTTTTATGCTTTTTTTTTTATTTTTTTTTTGGGGTAGTGTGAGT 1841

QY 587 CTCGCTTTGTCGATGATGGTACGCTGATACGTGCGAAATCTAAATGCTGATGAGTGG 646

Db 1842 ATGCTTTTTTGTGATGATGTACCTTTGGGAAATTAATTAATTTAAATTTAGTGTGTTT 1901

QY 647 CTCGTGTAATATTTGCTAAATAATTTGGAATATTCACACAGATTAATG 697

Db 1902 TAATTTTTTTATATTGTTTAAATTTGGAAATTTGGCATTCGATTTTATATGG 1952

```

RESULT 29
AB067197
ID AB067197 standard; DNA; 1163020 BP.
XX
XX AC AB067197;
XX
XX DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua contig DNA sequence #10.
XX
XX KW Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX LS Listeria innocua.
XX
XX PN WD200228891-A2.
XX
XX PD 11-APR-2002.
XX
XX PF 04-OCT-2001; 2001WO-FR03061.
XX
XX PR 04-OCT-2000; 2000FR-0012697.
XX
XX PS (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCT.
XX
XX PI Kunst F, Glaser P;
XX
XX DR WPI; 2002-332479/37.
XX
XX PT New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX
XX PS Claim 5; SEQ ID 10; 180bp; French.
XX
XX CC The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expressions. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX CC Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SO Sequence 1163020 BP; 388339 A; 197589 C; 235612 G; 341474 T; 6 other;
XX
XX Query Match 3.1%; Score 36.4; DB 24; Length 1163020;
XX Best Local Similarity 49.5%; Pred. No. 1.6e+02;
XX Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
XX
XX 36 TATTGCATTTGCTTATTCGATTCGGTTCGACATGTCGTAAGGTGAGCCGAAATTCGA 95
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 3116524 TGTTCGCTATATGTTTGAATTAATACAGCTATTTTGGTGTGCAAGGAAAAAGACGAGATGA 311713
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 96 ATGTGGACCACTTCATATACATCAATATTTTATACAGCTAATGCATTCGAAGGACATGT 155
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 311714	GCATTGGAACTTTTAAATGGAAGCGGATATTGAATGCTTCGTGAACATTTTAAACGAAGATGG	311773
Oy	156 TTATGTGAAGCTCTTTATGATCAAGAAGTGTCCTTAATGATGAAAGTGAGCGTCAATP	215
Db 311774	CCAAAGCGATTTTATATGACAGAACCAAGATTTTCACAAAGTACAGAAAGTTTGAAAGC	311833
Oy	216 TGCCGGCAATT	225
Db 311834	GGCTGGAAATT	311843

Query Match	3.1%	Score 36.4	DB 24	Length 3011208
Best Local Similarity	49.5%	Pred. NO.2.2e+02		
Matches 94	Conservative 0	Mismatches 96	Indels 0	Gaps 0
QY 36	TATTCGATTCCTTATTCGATTCGCGTTGACAAATGCTGCGAAGGTGAGCCGAAATTCGA	95		
DB 401918	TGTTCCGATATATGTTGATTAATACAGCTATTTTGGTGCAGGAAAGAAAGACGACGATGA	401977		
OY 96	ATGTGACCACTTCAATTAACATCAATTTTAATACAGTAAATGATTCGAAAGCACATCT	155		
DB 401978	GCATTTGCAACCTTTTAATGAAGCGGATGATGATGCTGTCGACATTTTAAGACGAAGATGG	402037		



OY 156 TTATGTGAAAGCTCTTTATGATCAAGAGTTCGCCATATGATGAGTGCAGCTCAGT 215  
 DB 402038 CCAAGCGATTATTATGACAGACAGAGATTCCACAAAGTACAGAGAGTTTGAAGC 402097  
 OY 216 TCCCGGAATT 225  
 DB 402098 GCGTGGAAATT 402107

## RESULT 31

ABA48451/C  
 ID ABA48451 standard; DNA: 510 BP.

AC ABA48451;

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #7146.

KW Human; microarray; single exon probe; gene expression; breast;  
 disease; cancer; ss.

OS Homo sapiens.

PN WO200157271-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,  
 useful for measuring gene expression in sample derived from human  
 breast, comprises number of single exon nucleic acid probes -

Claim 4; SEQ ID NO 7146; 327pp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon  
 nucleic acid probes for measuring gene expression in a sample derived  
 from human breast and BT 474 cells. The method involves contacting  
 the probes with a collection of detectably labelled nucleic acids  
 derived from mRNA of human breast, and then measuring the label  
 bound to each probe of the microarray. The probes are useful for  
 verifying the expression of regions of genomic DNA predicted to  
 encode proteins. They are useful for gene discovery, and for  
 determining predisposition and/or prognosing breast disease. Gene  
 expression analysis is useful for assessing the toxicity of chemical  
 agents on cells. The microarray of this invention presents a far greater  
 diversity of probes for measuring gene expression, with far less bias  
 than expressed sequence tag microarrays. The method is suitable for  
 rapid production of functional information from genomic sequence. The  
 present sequence is a single exon nucleic acid probe of the invention.  
 Note: The sequence data for this patent did not form part of the  
 printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 510 BP; 149 A; 184 C; 68 G; 109 T; 0 other;

Query Match

3.1%; Score 36.2; DB 22; Length 510;

Best Local Similarity 46.3%; Pred. No. 5.7;  
 Matches 119; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

OY 476 ATGGTGGACCAACCGGTCAACAGTTTCATTTGCTATGTCAGCCAGTTTATCAT 535  
 DB 414 ATGCTGATGACAGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 355  
 OY 536 AATGACATGCGATTCCTGTAACCGTTGATGATCTTTCGCGGTTGTCATTCCTGCTTG 595  
 DB 354 ATGTTGAAGCAGGTGCTGTAAGAGGTGATGATGATGATGATGATGATGATGATG 295  
 OY 596 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 655  
 DB 294 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 235  
 OY 656 AATATTTGCTAATATTAATTTGCAATATCCACAGATTTAATGCTGCCAAGAGCTCAG 715  
 DB 234 CAGGTGCTGTAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 175  
 OY 716 TATACAAATATGCGGAT 732  
 DB 174 CTGTAATGATGATGATGATGAT 158

## RESULT 32

ABA66355/C  
 ID ABA66355 standard; DNA: 510 BP.

AC ABA66355;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #14660.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for  
 analyzing gene expression in human foetal liver -

Claim 4; SEQ ID NO 14660; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for  
 measuring human gene expression in a sample derived from human foetal  
 liver. The single exon nucleic acid probes may be used for predicting,  
 measuring and displaying gene expression in samples derived from human  
 foetal liver. The present sequence is a single exon nucleic acid  
 probe of the invention.  
 Note: The sequence data for this patent did not form part of the  
 printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

CC



Query Match	Best Local Similarity	Score	DB	Length
Matches 119; Conservative	46.3%;	3.1%;	DB 22;	510;
	0;	Pred. No. 5.7;	Mismatches 138;	Indels 0; Gaps 0;
Sequence 510 BP; 149 A; 184 C; 68 G; 109 T; 0 other;				
416 ATGGTGACCAACCGCGTCAACCAAGTTCAATTTGCTATCATTTGGTCAGCCAGTTATCATATA	535			
414 ATGCGATGACGAGCGCCCTGATGCTGATGACAGATATTGTTGGTGCACAGTTGATGACG	355			
536 AATGACATCGCATCTCTAAACCCCTTGATACCTTTCTGGCGGTTTCCATTCTCTCTTG	595			
354 ATGTGGAAGCAGCGTCTGAAAAAGGTGTGATGACGAGGCGCAATCTGATGCTGTAGT	295			
536 TCGATGATGCTGACGCTGATACCTGGAATAATCTAAATGCTGATGATGCTGCTTGTATA	655			
294 ATGCGATGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	235			
656 AATATTTTCTTAAATATTTTGGAAATATCCACAGATTTAATGCTGCGCCAGCAAGCTCACG	715			
234 CAGGTGCTGAAAAATGCTGTTGATGCAAGTGCAGATGCTGATGCTGATGCTGATGCTGATG	175			
716 TATCAATATATGCGGAT	732			
174 CTGTAGTATGCTGCT	158			

Query Match	Best Local Similarity	Score	DB	Length
Matches 119;	Conservative 0;	Mismatches 138;	Indels 0;	Gaps 0;
476	ATGGTGGACCAACCGGTCCACAGTTCAATTTGCTATCATTTGGTGCAGCCAGTTTATCA	535		
414	ATGCTGATGCACAGCGCTGATGCTGATGCAGATGATTTGGTTGGTGCAGATGTTATGCAG	355		
536	AATGCACATGCGATTTCTGAACCGTTGATTTCTTTCGCGGGTTGCTCATTCCTCTTTCG	595		
354	ATGTTGAAACAGAGTGTGTAAGAAAGGTTGATGCAGAGTGCAGATGCTGATGCTGTAAGT	295		
596	TGAT	655		
294	ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	235		
656	AATATTTGCTAAT	715		
234	CAGGTGCTGTAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	175		
716	TATACAAATATTCGGAT 732			
174	CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	158		

RESULT 36  
 AAI46546/C  
 ID AAI46546 standard; DNA; 510 BP.  
 AC AAI46546;  
 DT 17-OCT-2001 (first entry)  
 DE Probe #15332 used to measure gene expression in human placenta sample.  
 EX Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; ss.  
 OS Homo sapiens.  
 PN WO200157272-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00663.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 WP: 2001-488897/53.  
 XR

Query	Match	Best Local Similarity	Score	DB	Length	Matches	119;	Conservative	0;	Mismatches	138;	Indels	0;	Gaps
XX	Human genome-derived single exon nucleic acid probes useful for													
PT	analyzing gene expression in human placenta -													
XX	Claim 25; SEQ ID No 15232; 654pp; English.													
PS														
XX	The present invention relates to single exon nucleic acid probes (SEN													
CC	CC The present sequence is one such probe. The probes are useful for													
CC	producing a microarray for predicting, measuring and displaying gene													
CC	expression in samples derived from human placenta. The probes are use													
CC	for antenatal diagnosis of human genetic disorders.													
XX														
XX	Sequence 510 BP; 149 A; 184 C; 68 G; 109 T; 0 other;													
SQL														
Query Match		3.1%;	Score 36.2;	DB 22;	Length 510;									
Best Local Similarity	46.3%;	Pred. No. 5.7;												
Matches 119;	Conservative	0;	Mismatches	138;	Indels	0;	Gaps							
QY	476 ATGTTGGACCAACCGGTCACCAAGTTCAATTGCTATCATTTGGTCAGCCAGTTATCATTA	535												
Db	414 ATGCTGATGACGAGGGCGTGAATGCTGATACAGGTATTTGGTGTGGTCAGATGTTGATGACAG	355												
QY	536 AATGCAATGCGATTTGGAACCGTTGATATCTTCTCGCGGTTGTCATCTCGCTTG	595												
Db	354 ATGTTGAAGACGAGTCTGCTGAAAAGGTTCATGACAGCTGCAGATCTGATCTGTGATG	295												
QY	596 TCGATGATGTTACGCGTGAATCTGTGGAATTTCTAAATGCTGATGATGATGCTCTTGATA	655												
Db	294 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	235												
QY	656 AATTTTGTCTAAATATTGGAATATCCAAACAGATTAAATGCTGCGCCAAAGACTCAGC	715												
Db	234 CAGCTGCTGAAAATGTTGATGATGCAAGTGCAGATGCTGATGCTGATGATGATGATGATG	175												
QY	716 TATCAATATATGCGGAT	732												
Db	174 CTGCTAGTATGATCTGCT	158												
RESULT 37														
ABK77860														
ID	ABK77860 standard; DNA; 1103 BP.													
AC	ABK77860;													
DT	13-AUG-2002 (first entry)													
XX	Bacillus clausii genomic sequence tag (GST) #703.													
DE	Differential gene expression; genomic sequenced tag; GST;													
KW	altered culture condition; environmental stress;													
KW	physiological provocation; ds.													
XX	Bacillus clausii.													
OS														
XX	W0200229113-A2.													
PN														
XX	11-APR-2002.													
PD														
XX	05-OCT-2001; 2001WO-US31437.													
PF														
XX	06-OCT-2000; 2000US-0680598.													
PR	27-MAR-2001; 2001US-279526P.													
XX	(NOVO ) NOVOZYMES BIOTECH INC.													
PA	(NOVO ) NOVOZYMES AS.													
XX	Berka R, Clausen IG;													

PT	Bacillus cells, by using substrate containing Bacillus genomic
PS	sequenced tag array -
XX	Claim 11; SEQ ID NO 5151; 200pp; English.
CC	The invention describes a method of monitoring differential expression of
CC	genes in a first Bacillus cell relative to expression of the genes in
CC	other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC	isolated from Bacillus cells to a substrate containing array of Bacillus
CC	genomic sequenced tags (GST), examining the array, and determining
CC	a relative gene expression by an observed hybridisation reporter signal of
CC	a spot in the array. The method is useful for measuring the expression of
CC	genes in a first Bacillus cell relative to expression of the same genes
CC	in one or more second Bacillus cells. The method is useful for monitoring
CC	global expression of several genes from a Bacillus cell, discovering new
CC	genes, identifying possible functions of unknown open reading frames and
CC	monitoring gene copy number variation and stability. Monitoring changes
CC	in expression of genes may be used to provide a representation of the way
CC	in which Bacillus cells adapt to changes in culture conditions,
CC	environmental stress or other physiological provocation. Extensive
CC	follow-up characterisation is unnecessary, when one spot on an array
CC	equals one gene or one open reading frame, since sequence information is
CC	available. This sequence represents a genomic sequence tag (GST) used in
CC	the method of the invention.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at
CC	ftp.wipo.int/pub/published_pct_sequences.
SQ	Sequence 1103 BP; 232 A; 272 C; 270 G; 329 T; 0 other;
OY	Query Match 3.1%; Score 36.2; DB 24; Length 1103;
Db	Best Local Similarity 56.2%; Pred. No. 8.1;
Matches	68; Conservative 0; Mismatches 53; Indels 0; Gaps 0
OY	471 TTTCGATGCTGGACCAACCAGTACCGTTCAATTGCTATCATTTGGTCACGCATTGA 530
Db	
OY	812 TTTTCAGGATGTGCTTGACGCTTACTTTCGTCGCGGTTGTCATTCCTCG 871
Db	
OY	531 TCATAAATGATGCGCATTCGTGAAACCGTTGTAATCTTCTGCGGCGTTCATTCCTCG 590
Db	
OY	872 TCATGCCGCGATTTGCAATTTTATCCTCTGCTTTTCCGCCCTGTGTCATCCGACCTG 931
OY	591 C 591
Db	
Db	932 C 932
RESULT 38	
AAS45434	
ID	AAS45434 standard; DNA: 7810 BP.
XX	AAS45434:
XX	18-DEC-2001 (first entry)
DE	Chemically pretreated genomic DNA associated with cell cycle #70.
XX	
KW	Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
KW	human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW	graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
KW	arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
KW	immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
KW	PCR primer.
OS	Homo sapiens.
PN	WO200168911-A2.
PD	20-SEP-2001.
PF	15-MAR-2001; 2001MO-EP02945.
XX	

PR 15-MAR-2000; 2000DE-1013847.  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPiG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI: 2001-602751/68.  
 XX  
 PT Designing primers and probes for analysing diseases associated with  
 PT cytosine methylation state e.g. arthritis, cancer, aging,  
 PT arteriosclerosis comprising fragments of chemically modified genes  
 PT associated with cell cycle -  
 XX  
 PS Claim 1; SEQ ID No 139; 28pp; English.  
 XX  
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA  
 CC molecules associated with the cell cycle and specific PCR primers of the  
 CC invention. The sequences are useful for detecting the methylation state  
 CC of all CpG dinucleotides in a sequence and therefore for analysing  
 CC associated diseases. By analysing cytosine methylations in the pretreated  
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy  
 CC of existing diseases or the predisposition to specific diseases can be  
 CC ascertained. The parameters may be compared to another set of genetic  
 CC and/or epigenetic parameters, the differences serving as basis for  
 CC diagnosis and/or prognosis events which are disadvantageous to patients.  
 CC The sequences of the invention are useful for the diagnosis and therapy  
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,  
 CC aging, glomerular disease, Lewy body disease, arthritis,  
 CC arteriosclerosis, solid tumours and cancers.  
 CC  
 SQ Sequence 7810 BP; 1928 A; 314 C; 2058 G; 3510 T; 0 other;  
 XX  
 Query Match 3.1%; Score 36.2; DB 22; Length 7810;  
 Best Local Similarity 59.0%; Pred. No. 20;  
 Matches 62; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
 OY 235 TTGATTCATGCAATGTCGCGGTACAGATCTCGAATCCACGNGATTTTGTATACA 294  
 DB 6550 TATGTTTAGTTGATGTTTCGTCGTAATTAATCTTTGGAGTGTGGAATGTTTATATT 6609  
 OY 235 ACACTGTTGTCATTTGTTTCATCCATTATTTGTACCAAGTT 339  
 DB 6610 TAGACGTTTGTGATGATTTTATAGATTATTTGATAAAGATT 6654  
 XX  
 RESULT 39  
 ABR28281  
 ID ABR28281 standard; DNA: 7810 BP.  
 XX  
 AC ABR28281;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE DNA transcription associated genomic DNA #78.  
 XX  
 KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;  
 KW PNA: cytosine methylation state; SNP; retroviral infection; gene; ds;  
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;  
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;  
 KW immunological disorder; Werner's syndrome; developmental disorder;  
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;  
 KW neurodegenerative disorder; Meanderburg syndrome; Niemann-Pick disease;  
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;  
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;  
 KW polyglutamine disorder; solid tumour.  
 XX  
 XX Unidentified.  
 XX  
 OS WO200192565-A2.  
 PN

XX  
 PD 06-DEC-2001.  
 XX  
 XX 06-APR-2001; 2001WO-EP03973.  
 XX  
 PF  
 XX  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPiG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI: 2002-090046/12.  
 XX  
 PT New nucleic acids or oligomers, useful for diagnosing or treating  
 PT diseases associated with DNA transcription, e.g. immunological  
 PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid  
 PT tumours or cancer -  
 XX  
 PS Claim 1; SEQ ID No 155; 32pp; English.  
 XX  
 CC The invention relates to a nucleic acid, which comprises a segment of the  
 CC chemically pretreated DNA of genes associated with DNA transcription from  
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide  
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical  
 CC to the chemically pretreated DNA of genes associated with DNA  
 CC transcription. The set of oligomer probes are useful for detecting the  
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)  
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for  
 CC diagnosing or treating diseases associated with DNA transcription  
 CC (particularly with the methylation status), e.g. adenosine deaminase  
 CC deficiency, viral infection, retroviral infection, Werner syndrome,  
 CC haematological disorders, immunological disorders, Werner syndrome,  
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,  
 CC neurological disorders, neurodegenerative disorders, Meanderburg  
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial  
 CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart  
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours  
 CC or cancer. Sequences ABR28127-ABR28472 represent DNA transcription  
 CC associated genomic DNA molecules of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from the  
 CC European Patent Office.  
 CC  
 SQ Sequence 7810 BP; 1928 A; 314 C; 2058 G; 3510 T; 0 other;  
 XX  
 Query Match 3.1%; Score 36.2; DB 24; Length 7810;  
 Best Local Similarity 59.0%; Pred. No. 20;  
 Matches 62; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
 OY 235 TTGATTCATGCAATGTCGCGGTACAGATCTCGAATCCACGNGATTTTGTATACA 294  
 DB 6550 TATGTTTAGTTGATGTTTCGTCGTAATTAATCTTTGGAGTGTGGAATGTTTATATT 6609  
 OY 235 ACACTGTTGTCATTTGTTTCATCCATTATTTGTACCAAGTT 339  
 DB 6610 TAGACGTTTGTGATGATTTTATAGATTATTTGATAAAGATT 6654  
 XX  
 RESULT 40  
 AAH34068/C  
 ID AAH34068 standard; cDNA: 1481 BP.  
 XX  
 AC AAH34068;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:1150.  
 XX  
 KW Human colon cancer; colon cancer antigen; diagnosis; detection;  
 KW colorectal carcinoma; ss.

XX	Homo sapiens.
XX	
XX	WO200122920-A2.
XX	
XX	05-APR-2001.
XX	
XX	28-SEP-2000; 2000WO-US26524.
XX	
XX	29-SEP-1999; 99US-0157137.
XX	03-NOV-1999; 99US-0163280.
XX	
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;
XX	
XX	WP1; 2001-235357/24.
DR	P-PSDB; AAG74663.
XX	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX	useful for preventing, diagnosing and/or treating colorectal cancers -
XX	
XX	Claim 1; Page 2962-2963; 9803pp; English.
XX	
XX	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where
CC	the proteins are collectively known as colon cancer antigens. The colon
CC	cancer antigens have cytostatic activity and can be used in gene
CC	therapy and vaccine production. N and P may be used in the prevention,
CC	diagnosis and treatment of diseases associated with inappropriate P
CC	expression. For example, N and P may be used to treat disorders
CC	associated with decreased expression by rectifying mutations or deletions
CC	in a patient's genome that affect the activity of P by expressing
CC	inactive proteins or to supplement the patient's own production of P.
CC	Additionally, N may be used to produce the colon cancer-associated ps,
CC	by inserting the nucleic acids into a host cell and culturing the cell
CC	to express the proteins. N and P can be used in the prevention, diagnosis
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC	present invention.
CC	
CC	N.B. Pages 666 to 662 and page 7053 of the sequence listing were
CC	missing at time of publication, meaning no sequences are present for
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 15:29:13 ; Search time 72 Seconds  
(without alignments)  
4945.163 Million cell updates/sec

Title: US-10-054-562a-3

Perfect score: 1161

Sequence: 1 atgatgattcgcctattatgc.....ttcgtccaatcagaagca 1161

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/6C.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/6D.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1161	100.0	1161	4	US-09-323-427-3
2	1161	100.0	1161	4	US-09-323-427-5
3	1161	100.0	1161	4	US-09-812-642-3
4	1161	100.0	1161	4	US-09-812-642-5
5	1161	100.0	1161	4	US-09-323-427-1
6	1161	100.0	1161	4	US-09-323-427-2
7	1161	100.0	1161	4	US-09-812-642-1
8	1161	100.0	1161	4	US-09-812-642-2
9	504.2	43.4	1372	4	US-09-323-427-6
10	504.2	43.4	1372	4	US-09-323-427-7
11	504.2	43.4	1372	4	US-09-812-642-6
12	504.2	43.4	1372	4	US-09-812-642-7
13	461	39.7	892	4	US-09-323-427-16
14	461	39.7	892	4	US-09-323-427-17
15	461	39.7	892	4	US-09-812-642-16
16	461	39.7	892	4	US-09-812-642-18
17	310.4	26.7	813	4	US-09-323-427-8
18	310.4	26.7	813	4	US-09-323-427-10
19	310.4	26.7	813	4	US-09-812-642-8
20	310.4	26.7	813	4	US-09-812-642-10
21	35.4	3.0	348	4	US-09-308-003-45
22	35.4	3.0	1182	4	US-09-308-003-8
23	34.2	2.9	4865	3	US-08-894-017-24
24	34	2.9	34	4	US-09-323-427-11
25	34	2.9	34	4	US-09-323-427-12
26	34	2.9	34	4	US-09-812-642-11
27	34	2.9	34	4	US-09-812-642-12

28	34	2.9	895	4	US-09-602-877A-91	Sequence 91, Appl
29	34	2.9	1521	4	US-09-134-001C-1445	Sequence 1445, Ap
30	33.6	2.9	2427	4	US-09-134-001C-268	Sequence 268, App
31	33.4	2.9	1867	1	US-07-772-087-1	Sequence 1, Appl
32	33.4	2.9	16950	4	US-09-453-702B-166	Sequence 166, App
33	33	2.8	11225	6	5182210-9	Patent No. 5182210
34	32.6	2.8	630	4	US-09-385-982-45	Sequence 45, Appl
35	32.6	2.8	5177	6	5352450-1	Patent No. 5352450
36	32.4	2.8	576	5	PCT-US96-05320A-705	Sequence 705, App
37	32.4	2.8	4090	2	US-08-781-802-5	Sequence 5, Appl
38	32.4	2.8	4090	4	US-08-694-078-5	Sequence 5, Appl
39	32.4	2.8	31960	4	US-09-453-702B-11	Sequence 11, Appl
40	32.2	2.8	1860	4	US-09-134-001C-1411	Sequence 1411, Ap
41	32.2	2.8	2121	4	US-09-134-001C-2071	Sequence 2071, Ap
42	32	2.8	1709	4	US-08-936-165A-188	Sequence 188, App
43	32	2.8	2338	1	US-08-425-069-1	Sequence 1, Appl
44	32	2.8	2338	2	US-08-317-844B-1	Sequence 1, Appl
45	31.8	2.7	12666	4	US-08-961-527-137	Sequence 137, App

#### ALIGNMENTS

```
RESULT 1
US-09-323-427-3
; Sequence 3, Application US/09323427
; Patent No. 6248329
; GENERAL INFORMATION:
; APPLICANT: Chandrasekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid
; FILE REFERENCE: HW-8
; CURRENT APPLICATION NUMBER: US/09/323.427
; EARLIER FILING DATE: 1999-06-01
; EARLIER APPLICATION NUMBER: 60/087,435
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Dirofilaria immitis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1161)
US-09-323-427-3
Query Match      100.0%; Score 1161; DB 4; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGATTCGTCCTTATGCTTCTGACCTTATGTCATTTGTCATTCGTCGCG 60
    |||||||
Db 1 ATGATGATTCGTCCTTATGCTTCTGACCTTATGTCATTTGTCATTCGTCGCG 60

QY 61 GTTGACATGTCGTGTCGAGTGACGACGAATGAGTGACCACTTCATTAATCAATC 120
    |||||||
Db 1 GTTGACATGTCGTGTCGAGTGACGACGAATGAGTGACCACTTCATTAATCAATC 120

QY 121 AATTTTAAATACAGTAATGATGATTCGAGGAGCATTTATGGAAGGCTTTATGATCAA 180
    |||||||
Db 1 AATTTTAAATACAGTAATGATGATTCGAGGAGCATTTATGGAAGGCTTTATGATCAA 180

QY 181 GAAGTTGCGGTAAATGATGAGAGTGACGTCACGTCGCGGAATTCACCTTCATTTGAT 240
    |||||||
Db 181 GAAGTTGCGGTAAATGATGAGAGTGACGTCACGTCGCGGAATTCACCTTCATTTGAT 240

QY 241 TCATGCAATGTCGTCGTCATGACGATCTCTGATTCAGATGATTTTGAACAACACT 300
    |||||||
Db 241 TCATGCAATGTCGTCGTCATGACGATCTCTGATTCAGATGATTTTGAACAACACT 300

QY 301 GTTGTCATTCGTCATTCATTCATTTGTTATCAACAAAGTTGATTCGTCATTCAGTACAA 360
    |||||||
```

```

Db 301 GTTGCATTCGTTTCATTCATTAATTTGTTACCAAGTGAATCGTGCAATTCGATACAA 360
Oy 361 TGCCTTTACATGAGAGCTGATTAACACATAGTAGTGACAGATTGAGGTATCTGAATACCA 420
Db 361 TGCCTTTACATGAGAGCTGATTAACACATAGTAGTGACAGATTGAGGTATCTGAATACCA 420
Oy 421 ACTGCTTTCAACATCAATATGTCGCCGATGCGAGTATCGCTTATGAATTTGGATGATG 480
Db 421 ACTGCTTTCAACATCAATATGTCGCCGATGCGAGTATCGCTTATGAATTTGGATGATG 480
Oy 481 GGACCAACCGGTGACACAGTTCAATTTGCTATCATTTGGTCAAGCAGTTTATCATTAATG 540
Db 481 GGACCAACCGGTGACACAGTTCAATTTGCTATCATTTGGTCAAGCAGTTTATCATTAATG 540
Oy 541 ACATGCGATTCTGAACCGTGTGATCTTCTGCGCGGTGTCATTCCTGCTTTGTGAT 600
Db 541 ACATGCGATTCTGAACCGTGTGATCTTCTGCGCGGTGTCATTCCTGCTTTGTGAT 600
Oy 601 GATGCTAACCGGTGATCTGTTGGAATTTCTAATGCTGATGATGCTGCTTATGAATAT 660
Db 601 GATGCTAACCGGTGATCTGTTGGAATTTCTAATGCTGATGATGCTGCTTATGAATAT 660
Oy 661 TTGCTAATTAATTTGGAATTCACACAGATTAAATGCGTGGCCAAAGCTCACGTATAC 720
Db 661 TTGCTAATTAATTTGGAATTCACACAGATTAAATGCGTGGCCAAAGCTCACGTATAC 720
Oy 721 AAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 AAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Oy 781 CCAATATGCGAATGTTTTCACACAGATTGTTCAAGAACCAAGGATTCGAGCTGTTAAA 840
Db 781 CCAATATGCGAATGTTTTCACACAGATTGTTCAAGAACCAAGGATTCGAGCTGTTAAA 840
Oy 841 ACAGGTGCTGCGGACGACAAACCTGCTGACAGTGGCGCAACTGCTTACTCAAGAAAAGA 900
Db 841 ACAGGTGCTGCGGACGACAAACCTGCTGACAGTGGCGCAACTGCTTACTCAAGAAAAGA 900
Oy 901 TCGCAGAACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATG 960
Db 901 TCGCAGAACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATG 960
Oy 961 GATGATATCAACCTTTCGACAGTTGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 GATGATATCAACCTTTCGACAGTTGATGATGATGATGATGATGATGATGATGATGATG 1020
Oy 1021 CAACCTGTAATCTTCTGACAGTAAATGGAATCTGATGTCACCATTTGGCTTCTCA 1080
Db 1021 CAACCTGTAATCTTCTGACAGTAAATGGAATCTGATGTCACCATTTGGCTTCTCA 1080
Oy 1081 AGTTTATGAGGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 AGTTTATGAGGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Oy 1141 TTTTCGCAAAATCAGAAGCA 1161
Db 1141 TTTTCGCAAAATCAGAAGCA 1161

```

```

RESULT 2
US-09-323-427-5/c
; Sequence 5, Application US/09323427
; Patent No. 6248329
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: Parasitic Helminth Cuticulin Proteins, Nucleic Acid
; FILE REFERENCE: Hw-8
; CURRENT APPLICATION NUMBER: US/09/323,427
; CURRENT FILING DATE: 1999-06-01
; EARLIER APPLICATION NUMBER: 60/087,435

```

```

; EARLIER FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Dicrofilaria immitis
US-09-323-427-5

Query Match      100.0%; Score 1161; DB 4; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGATGATGCTTATGCTTATGCTTCTGATGATGATGATGATGATGATGATGATGATG 60
Db 1161 ATGATGATGCTTATGCTTATGCTTCTGATGATGATGATGATGATGATGATGATGATG 1102
Oy 61 GTTGACAAATGCTGTCGAAGGTGAGCCGGAATTTGAATGAGACCAACTTAATTAACAATC 120
Db 1101 GTTGACAAATGCTGTCGAAGGTGAGCCGGAATTTGAATGAGACCAACTTAATTAACAATC 1042
Oy 121 AAATTAATACAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 1041 AAATTAATACAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 982
Oy 181 GAAGGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 981 GAAGGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 922
Oy 241 TCATGCAATGTCGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 921 TCATGCAATGTCGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 862
Oy 301 GTTGCATTCGTTTCATCATTAATTTGTTTACCAAGTTGATGTCATATGAGTACAA 360
Db 861 GTTGCATTCGTTTCATCATTAATTTGTTTACCAAGTTGATGTCATATGAGTACAA 802
Oy 361 TGCCTTTACATGAGAGCTGATTAACAGATTAAGTATGATGATGATGATGATGATGATG 420
Db 801 TGCCTTTACATGAGAGCTGATTAACAGATTAAGTATGATGATGATGATGATGATGATG 742
Oy 421 ACTGCTTTCAACATCAATATGTCGCCGATGCGGATGATGATGATGATGATGATGATGAT 480
Db 741 ACTGCTTTCAACATCAATATGTCGCCGATGCGGATGATGATGATGATGATGATGATGAT 682
Oy 481 GGACCAACCGGTGACACAGTTCAATTTGCTATCATTTGTCACAGCAGTTTATCATTAATG 540
Db 681 GGACCAACCGGTGACACAGTTCAATTTGCTATCATTTGTCACAGCAGTTTATCATTAATG 622
Oy 541 ACATGCGATTCTGAACACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 621 ACATGCGATTCTGAACACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 562
Oy 601 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 561 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
Oy 661 TTGCTAATTAATTTGGAATATCAACAGATTTAATGGTGGCCAGAAAGCTCACGTATAC 720
Db 501 TTGCTAATTAATTTGGAATATCAACAGATTTAATGGTGGCCAGAAAGCTCACGTATAC 442
Oy 721 AAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 441 AAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 382
Oy 781 CCAATATGCGAATGTTGTCGACACCAATGTTTCAAGCAACCAAGATTTGGAGCTGTTAAA 840
Db 381 CCAATATGCGAATGTTGTCGACACCAATGTTTCAAGCAACCAAGATTTGGAGCTGTTAAA 322
Oy 841 ACAGGTGCTGCGGACGACAAACCTGCTGACAGTGGCGCAACTGCTTACTCAAGAAAAGA 900
Db 321 ACAGGTGCTGCGGACGACAAACCTGCTGACAGTGGCGCAACTGCTTACTCAAGAAAAGA 262

```



SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 1161  
 ; TYPE: DNA  
 ; ORGANISM: *Dirofilaria immitis*  
 US-09-812-642-5

Query Match 100.0%; Score 1161; DB 4; Length 1161;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATGATGCTCTTATGCTTCTGTAACACTTATGATGCTTATGATTCGG 60
    |||
DB 1161 ATGATGATGCTCTTATGCTTCTGTAACACTTATGATGCTTATGATTCGG 1102
QY 61 GTTGACATGGTGTGCAAGGTGACCCGAATTAATGTGACCACTTCATATCAATC 120
    |||
DB 1101 GTTGACATGGTGTGCAAGGTGACCCGAATTAATGTGACCACTTCATATCAATC 1042
QY 121 AATTTAATACAGTAATGATTCAGAGACATGTTATGGAAGGCTTTATGATCAA 180
    |||
DB 1041 AATTTAATACAGTAATGATTCAGAGACATGTTATGGAAGGCTTTATGATCAA 982
QY 181 GAAGGTGCGGTATGATGAGAGGTGACGTCAGTTGCCGGAATTTCACTTCATTTGAT 240
    |||
DB 981 GAAGGTGCGGTATGATGAGAGGTGACGTCAGTTGCCGGAATTTCACTTCATTTGAT 922
QY 241 TCATGCAATTTGCGGTACACGATCTCTGAATCCACGTGTATTTTGTAGCAACAAT 300
    |||
DB 921 TCATGCAATTTGCGGTACACGATCTCTGAATCCACGTGTATTTTGTAGCAACAAT 862
QY 301 GTTGCTATTTGCTTTCATTCATATTTGTACCAAGTTGATGTCATTCGATTCAGATCAA 360
    |||
DB 861 GTTGCTATTTGCTTTCATTCATATTTGTACCAAGTTGATGTCATTCGATTCAGATCAA 802
QY 361 TCGTTTACATGAGAGCTGATTAACAGAGTAGTGACAGATGAGATGTGAATACACA 420
    |||
DB 801 TCGTTTACATGAGAGCTGATTAACAGAGTAGTGACAGATGAGATGTGAATACACA 742
QY 421 ACTGCTTTTCAACTCAATGTGCCGATGCCGATGCCGATATGCAATTTTGGATGGT 480
    |||
DB 741 ACTGCTTTTCAACTCAATGTGCCGATGCCGATGCCGATATGCAATTTTGGATGGT 682
QY 481 GGACCAACGGGTCAACAGATTCATATGTCATGTCAGGCACTTTATATCAATATG 540
    |||
DB 681 GGACCAACGGGTCAACAGATTCATATGTCATGTCAGGCACTTTATATCAATATG 622
QY 541 ACATGCGATTTGAAACCGTTGATCTTCTGCGCGTTGTCATCTGCTTTGCGAT 600
    |||
DB 621 ACATGCGATTTGAAACCGTTGATCTTCTGCGCGTTGTCATCTGCTTTGCGAT 562
QY 601 GATGTAAGGATGATCTGTGAAATTTCTAATGCTGATGATGTGCTTGTATTAAT 660
    |||
DB 561 GATGTAAGGATGATCTGTGAAATTTCTAATGCTGATGATGTGCTTGTATTAAT 502
QY 661 TTGCTAATTAATTTGGAATATCCACAGATTAATGGTGGCAAGAACCTCAGATATAC 720
    |||
DB 501 TTGCTAATTAATTTGGAATATCCACAGATTAATGGTGGCAAGAACCTCAGATATAC 442
QY 721 AATATGCGGATGATCACAGCTTTTCTATCAATGCCAGATAGATTAATCAATTAAGAA 780
    |||
DB 441 AATATGCGGATGATCACAGCTTTTCTATCAATGCCAGATAGATTAATCAATTAAGAA 382
QY 781 CCAATATAGGATGTGTCGACACAAATGTTAAGAACCAAGAGATTGCGACCTGTAA 840
    |||
DB 381 CCAATATAGGATGTGTCGACACAAATGTTAAGAACCAAGAGATTGCGACCTGTAA 322
QY 841 ACAGGTGTCGCGAGCAAAACCTGTCGAGTGGCAACTGTTTCTCAAGAAAGA 900
    |||
DB 321 ACAGGTGTCGCGAGCAAAACCTGTCGAGTGGCAACTGTTTCTCAAGAAAGA 262
QY 901 TCTGAGAAACCGAGATATATCATTTAGTACAACTGATATCAACCCCTGAAATTAGC 960
    |||

```

```

DB 261 TCTGACAGACCGGAGATATCATTTGATGTACGAACCTGATATCAACCCCTGAAATTAGC 202
QY 961 GATGATATCAAGCTTTGCCAGTTGATTTACGTACCGGTGACTTGTCAACATATAGA 1020
    |||
DB 201 GATGATATCAAGCTTTGCCAGTTGATTTACGTACCGGTGACTTGTCAACATATAGA 142
QY 1021 CAACCTGTATACTTCTGTCGATRCAAAATGGAATGTGATGTGCACATTTGGCTTCYA 1080
    |||
DB 141 CAACCTGTATACTTCTGTCGATRCAAAATGGAATGTGATGTGCACATTTGGCTTCYA 82
QY 1081 ATGTTTATGAGTTTAACCATTCATGATGCTGCGCGTCATTTATACATTTGGTTAAA 1140
    |||
DB 81 ATGTTTATGAGTTTAACCATTCATGATGCTGCGCGTCATTTATACATTTGGTTAAA 22
QY 1141 TTTCGTCCAATCAGAAGGCA 1161
    |||
DB 21 TTTCGTCCAATCAGAAGGCA 1

```

RESULT 5.  
 US-09-323-427-1  
 ; Sequence 1, Application US/09323427  
 ; Patent No. 6248329  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chandrasekar, Ramaswamy  
 ; APPLICANT: Morales, Tony H.  
 ; TITLE OF INVENTION: Parasitic Helminth Cuticulin Proteins, Nucleic Acid  
 ; FILE REFERENCE: MW-8  
 ; CURRENT APPLICATION NUMBER: US/09/323,427  
 ; EARLIER FILING DATE: 1999-06-01  
 ; EARLIER APPLICATION NUMBER: 60/087,435  
 ; EARLIER FILING DATE: 1998-06-01  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1779  
 ; TYPE: DNA  
 ; ORGANISM: *Dirofilaria immitis*  
 US-09-323-427-1

Query Match 100.0%; Score 1161; DB 4; Length 1779;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATGATGCTCTTATGCTTCTGTAACACTTATGATGCTTATGATTCGG 60
    |||
DB 167 ATGATGATGCTCTTATGCTTCTGTAACACTTATGATGCTTATGATTCGG 226
QY 61 GTTGACATGGTGTGCAAGGTGACCCGAATTAATGTGACCACTTCATATCAATC 120
    |||
DB 227 GTTGACATGGTGTGCAAGGTGACCCGAATTAATGTGACCACTTCATATCAATC 286
QY 121 AATTTAATACAGTAATGATTCAGAGACATGTTATGGAAGGCTTTATGATCAA 180
    |||
DB 287 AATTTAATACAGTAATGATTCAGAGACATGTTATGGAAGGCTTTATGATCAA 346
QY 181 GAAGGTGCGGTATGATGAGAGGTGACGTCAGTTGCCGGAATTTCACTTCATTTGAT 240
    |||
DB 347 GAAGGTGCGGTATGATGAGAGGTGACGTCAGTTGCCGGAATTTCACTTCATTTGAT 406
QY 241 TCATGCAATTTGCGGTACACGATCTCTGAATCCACGCTGTATTTGTAAACAACAT 300
    |||
DB 407 TCATGCAATTTGCGGTACACGATCTCTGAATCCACGCTGTATTTGTAAACAACAT 466
QY 301 GTTGCTATTTGCTTTCATTCATATTTGTACCAAGTTGATGTCATTCGATTCAGATCAA 360
    |||
DB 467 GTTGCTATTTGCTTTCATTCATATTTGTACCAAGTTGATGTCATTCGATTCAGATCAA 526
QY 361 TCTGTTTACATGAGAGCTGATTAACAGATAGTACACAGATTTGAGATTTGGAATAGCA 420
    |||
DB 527 TCTGTTTACATGAGAGCTGATTAACAGATAGTACACAGATTTGAGATTTGGAATAGCA 586
    |||

```

```
QY 421 ACTGCTTTCAAACTCAAAATTTGCCGATGCCAGTATCCGTTATGAAATTTTGGATGT 480
|
|
|
Db 587 ACTGCTTTCAAACTCAAAATTTGCCGATGCCAGTATCCGTTATGAAATTTTGGATGT 646
|
|
|
QY 481 GGACCAACCGGTCACACAGTTCATTTTGTATCATTTGTCAGCCAGTTTATCATTAATGG 540
|
|
|
Db 647 GGACCAACCGGTCACACAGTTCATTTTGTATCATTTGTCAGCCAGTTTATCATTAATGG 706
|
|
|
QY 541 ACATGCGATTCGAAACCGTGTGATCTTTCGCGCGGTGTGCCATTCCTGCTGTGTGAT 600
|
|
|
Db 707 ACATGCGATTCGAAACCGTGTGATCTTTCGCGCGGTGTGCCATTCCTGCTGTGTGAT 766
|
|
|
QY 601 GATGTAACGGTGTGATCTGTAATTCCTAATGCTGATGATGATGCTGTGTGATAATAT 660
|
|
|
Db 767 GATGTAACGGTGTGATCTGTAATTCCTAATGCTGATGATGATGCTGTGTGATAATAT 826
|
|
|
QY 661 TTGCTAATTAATTTGGAATTCACACAGATTTAATGCTGTGCCAAGAGCTCACGTATAC 720
|
|
|
Db 827 TTGCTAATTAATTTGGAATTCACACAGATTTAATGCTGTGCCAAGAGCTCACGTATAC 886
|
|
|
QY 721 AAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
|
|
|
Db 887 AAATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 946
|
|
|
QY 781 CCAAAATAGCGAATGTGTTCACCAAAATGTTCAGAACCAAGAGATTCGAGCTGTAA 840
|
|
|
Db 947 CCAAAATAGCGAATGTGTTCACCAAAATGTTCAGAACCAAGAGATTCGAGCTGTAA 1006
|
|
|
QY 841 ACAGGTGTGCGCGACGACCAAACTGCTGACGCTGCCAATCTGCTTACTCAAGAAAAGA 900
|
|
|
Db 1007 ACAGGTGTGCGCGACGACCAAACTGCTGACGCTGCCAATCTGCTTACTCAAGAAAAGA 1066
|
|
|
QY 901 TCTGCACACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
|
|
|
Db 1067 TCTGCACACCGGAGATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
|
|
|
QY 961 GATGATATCAAGCTTTGCCAGTTGATTTAGCTCACCGTGCACCTTCTGCAACATTAATGA 1020
|
|
|
Db 1127 GATGATATCAAGCTTTGCCAGTTGATTTAGCTCACCGTGCACCTTCTGCAACATTAATGA 1186
|
|
|
QY 1021 CAACCTGTAATCTTGTCTGCGATGACAAATGGAATCTGCAATGATGATGATGATGATGATGAT 1080
|
|
|
Db 1187 CAACCTGTAATCTTGTCTGCGATGACAAATGGAATCTGCAATGATGATGATGATGATGATGAT 1246
|
|
|
QY 1081 ATGTTTATGCGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
|
|
|
Db 1247 ATGTTTATGCGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
|
|
|
QY 1141 TTTGCTCCAAATCAGAAAGCA 1161
|
|
|
Db 1307 TTTGCTCCAAATCAGAAAGCA 1327
|
|
|
RESULT 6
US-09-323-427-2/c
; Sequence 2, Application US/09323427
; Patent No. 6248329
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid
; FILE REFERENCE: MW-8
; CURRENT APPLICATION NUMBER: US/09/323,427
; CURRENT FILING DATE: 1999-06-01
; EARLIER APPLICATION NUMBER: 60/087,435
; EARLIER FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Dicrofilaria immitis
```

```
US-09-323-427-2
Query Match 100.0%; Score 1161; DB 4; Length 1779;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGATTCGCTTAAATGCTTTCTGTAATGCTGATGATGATGATGATGATGATGATGATGAT 60
|
|
|
Db 1613 ATGATGATTCGCTTAAATGCTTTCTGTAATGCTGATGATGATGATGATGATGATGATGATGAT 1554
|
|
|
QY 61 GTTGACAATGCTGTGCAAGGTGACCCAGAAATTAATGTCAGCACTTCAATTAATC 120
|
|
|
Db 1553 GTTGACAATGCTGTGCAAGGTGACCCAGAAATTAATGTCAGCACTTCAATTAATC 1494
|
|
|
QY 121 AATTTTAATNACGTAATGCTGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
|
|
|
Db 1493 AATTTTAATNACGTAATGCTGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
|
|
|
QY 181 GAAAGTTGCCGTAATGATGAAAGGTGACGTCAGATGTCGGAATTCACCTTCAATTTGAT 240
|
|
|
Db 1433 GAAAGTTGCCGTAATGATGAAAGGTGACGTCAGATGTCGGAATTCACCTTCAATTTGAT 1374
|
|
|
QY 241 TCATGCAATGTTGCGCGTACACGATCTCTGATTCAGTGTATTTTGTAAACAACACT 300
|
|
|
Db 1373 TCATGCAATGTTGCGCGTACACGATCTCTGATTCAGTGTATTTTGTAAACAACACT 1314
|
|
|
QY 301 GTTGCAATTTGCTTCAATCCATTTATTTGTACCAAGTTGATCGTATTCAGTATGATACAA 360
|
|
|
Db 1313 GTTGCAATTTGCTTCAATCCATTTATTTGTACCAAGTTGATCGTATTCAGTATGATACAA 1254
|
|
|
QY 361 TGTCTTTACATGGAAGCTGATTAACAGTATGATGTCACAGATTTGAGTATCTGAATTCACA 420
|
|
|
Db 1253 TGTCTTTACATGGAAGCTGATTAACAGTATGATGTCACAGATTTGAGTATCTGAATTCACA 1194
|
|
|
QY 421 ACTGCTTTCAAACTCAAAATTTGCCGATGCCAGTATCCGTTATGAAATTTTGGATGT 480
|
|
|
Db 1193 ACTGCTTTCAAACTCAAAATTTGCCGATGCCAGTATCCGTTATGAAATTTTGGATGT 1134
|
|
|
QY 481 GGACCAACCGGTCACACAGTTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
|
|
|
Db 1133 GGACCAACCGGTCACACAGTTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
|
|
|
QY 541 ACATGCGATTCGAAACCGTGTGATCTTTCGCGGGTGTGTCATTCCTGCTTGTGAT 600
|
|
|
Db 1073 ACATGCGATTCGAAACCGTGTGATCTTTCGCGGGTGTGTCATTCCTGCTTGTGAT 1014
|
|
|
QY 601 GATGTAACGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
|
|
|
Db 1013 GATGTAACGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
|
|
|
QY 661 TTTGCTAATTAATTTGGAATTCACACAGATTTAATGCTGTGCCAAGAGCTCACGTATAC 720
|
|
|
Db 953 TTTGCTAATTAATTTGGAATTCACACAGATTTAATGCTGTGCCAAGAGCTCACGTATAC 894
|
|
|
QY 721 AAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
|
|
|
Db 893 AAATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
|
|
|
QY 781 CCAAAATAGCGAATGTGTTCGACCAATGTCACACAGATTTGGAAGCTGTAA 840
|
|
|
Db 833 CCAAAATAGCGAATGTGTTCGACCAATGTCACACAGATTTGGAAGCTGTAA 774
|
|
|
QY 841 ACAGGTGTGCGCGACGACCAAACTGCTGACGCTGCCAATCTGCTTACTCAAGAAAAGA 900
|
|
|
Db 773 ACAGGTGTGCGCGACGACCAAACTGCTGACGCTGCCAATCTGCTTACTCAAGAAAAGA 714
|
|
|
QY 901 TCTGCAGAACCGGAGAAATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
|
|
|
Db 713 TCTGCAGAACCGGAGAAATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
|
|
|
QY 961 GATGATATCAAGCTTTGCCAGTTGATTTAGCTCACCGTGCACCTTCTGCAACATTAATGA 1020
|
|
|
Db 653 GATGATATCAAGCTTTGCCAGTTGATTTAGCTCACCGTGCACCTTCTGCAACATTAATGA 594
|
|
|
```

QY 1021 CAACCTGTAATCTGCTGACGATCAAAATGGAATCTGCATGCATCCATTTGGCTTCGA 1080  
 |||||||  
 Db 593 CAACCTGTAATCTGCTGACGATCAAAATGGAATCTGCATGCATCCATTTGGCTTCGA 534  
 |||||||  
 QY 1081 ATGTTATGAGGTTTAAGCATTTGATTTGCTGCCGTCATTTATTTACCTTTGGTTTAA 1140  
 |||||||  
 Db 533 ATGTTATGAGGTTTAAGCATTTGATTTGCTGCCGTCATTTATTTACCTTTGGTTTAA 474  
 |||||||  
 QY 1141 TTTCGTCCAATCAGAAGCA 1161  
 |||||||  
 Db 473 TTTCGTCCAATCAGAAGCA 453  
 |||||||

# RESULT 7 US-09-812-642-1

; Sequence 1, Application US/09812642  
 ; Patent No. 6368600  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chandrasekar, Ramaswamy  
 ; APPLICANT: Morales, Tony H.  
 ; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
 ; FILE REFERENCE: HW-8  
 ; CURRENT APPLICATION NUMBER: US/09/812,642  
 ; CURRENT FILING DATE: 2001-03-20  
 ; PRIOR APPLICATION NUMBER: 09/323,427  
 ; PRIOR FILING DATE: 1999-06-01  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: Patentln Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1779  
 ; TYPE: DNA  
 ; ORGANISM: Dirofilaria immitis  
 ; US-09-812-642-1

Query Match 100.0%; Score 1161; DB 4; Length 1779;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATGATTCGCTTATTTGCTTTCTGCTACTACATTTATTCATTTGCTTATTCGATTCG 60  
 |||||||  
 Db 167 ATGATGATTCGCTTATTTGCTTTCTGCTACTACATTTATTCATTTGCTTATTCGATTCG 226  
 |||||||  
 QY 61 GTTGACAATGCTGTCGAGAGTGGCCAGAAATTTGAATGTCGACCACTTCATTAACATC 120  
 |||||||  
 Db 227 GTTGACAATGCTGTCGAGAGTGGCCAGAAATTTGAATGTCGACCACTTCATTAACATC 286  
 |||||||  
 QY 121 AATTTAATACAGTAATGCATTCGAAAGACATGTTATGTAAGTCTTTATGATCA 180  
 |||||||  
 Db 287 AATTTAATACAGTAATGCATTCGAAAGACATGTTATGTAAGTCTTTATGATCA 346  
 |||||||  
 QY 181 GAAGGTGCCGTAATGATGAGAGTGCAGTCAAGTTCGCGAATTTCACTTCATTTGAT 240  
 |||||||  
 Db 347 GAAGGTGCCGTAATGATGAGAGTGCAGTCAAGTTCGCGAATTTCACTTCATTTGAT 406  
 |||||||  
 QY 241 TCATGCAATGTCGCGTACAGATCTCGAATCCACGTCGATTTTGTATTTGTAACAACACT 300  
 |||||||  
 Db 407 TCATGCAATGTCGCGTACAGATCTCGAATCCACGTCGATTTTGTATTTGTAACAACACT 466  
 |||||||  
 QY 301 GTTGCAATTTGCTTATTCATTTATTTGTTACCAAGTTGATCGGATTCGATTCAGATCA 360  
 |||||||  
 Db 467 GTTGCAATTTGCTTATTCATTTATTTGTTACCAAGTTGATCGGATTCGATTCAGATCA 526  
 |||||||  
 QY 361 TGCCTTTACATGAGAGTCGATTAACAGTATGTCACAGATTCGATTCGATTCAGATCA 420  
 |||||||  
 Db 527 TGCCTTTACATGAGAGTCGATTAACAGTATGTCACAGATTCGATTCGATTCAGATCA 586  
 |||||||  
 QY 421 ACTGCTTTTCAACGCAATGTCGCGATTCGACAGATTCGCTTATGAAATTTGGATGT 480  
 |||||||  
 Db 587 ACTGCTTTTCAACGCAATGTCGCGATTCGACAGATTCGCTTATGAAATTTGGATGT 646  
 |||||||  
 QY 481 GGACCAACCGGTCAACGATTCATTTGCTATTCATTTGCTACGCAAGTTATCATTAATAG 540  
 |||||||

|||||  
 Db 647 GGACCAACCGGTCAACGATTCATTTGCTATTCATTTGCTACGAGATTCATTAATAG 706  
 |||||||  
 QY 541 ACATGCGATTCGAAACCGTTCATTTGCTGCGCGGTTCGATTCATTCCTGCTTGTGAT 600  
 |||||||  
 Db 707 ACATGCGATTCGAAACCGTTCATTTGCTGCGCGGTTCGATTCATTCCTGCTTGTGAT 766  
 |||||||  
 QY 601 GATGTAACCGGTTCATTTGCTGAAATTCGAAATCTGATGATTCGCTTGTGATTAAT 660  
 |||||||  
 Db 767 GATGTAACCGGTTCATTTGCTGAAATTCGAAATCTGATGATTCGCTTGTGATTAAT 826  
 |||||||  
 QY 661 TTGCTAAATTAATTTGGAATTCACACAGATTTAATGCTGCGCAAGAGTCACGATTA 720  
 |||||||  
 Db 827 TTGCTAAATTAATTTGGAATTCACACAGATTTAATGCTGCGCAAGAGTCACGATTA 886  
 |||||||  
 QY 721 AATATGCGGATGCATACAGCTTTTCTATCATTCAGATTCAGATTAATTAAGAA 780  
 |||||||  
 Db 887 AATATGCGGATGCATACAGCTTTTCTATCATTCAGATTCAGATTAATTAAGAA 946  
 |||||||  
 QY 781 CCAATATGCGAATGTTTCGACCACAATGTTTCAGAACCAAGATTCGAGCTTTAA 840  
 |||||||  
 Db 947 CCAATATGCGAATGTTTCGACCACAATGTTTCAGAACCAAGATTCGAGCTTTAA 1006  
 |||||||  
 QY 841 ACAAGTGTGTCGCCAGCAAACTGCTGACGTCGCGCACTTCGTTACTCAAGAAAGA 900  
 |||||||  
 Db 1007 ACAAGTGTGTCGCCAGCAAACTGCTGACGTCGCGCACTTCGTTACTCAAGAAAGA 1066  
 |||||||  
 QY 901 TCTGCAACCGGAGATATCATTTGATGTCAGAACATGATTCACACCTTCGAAATAGC 960  
 |||||||  
 Db 1067 TCTGCAACCGGAGATATCATTTGATGTCAGAACATGATTCACACCTTCGAAATAGC 1126  
 |||||||  
 QY 961 GATGATTAATCAACCTTTGCGAGTTGATTAACGTCACCGTCACCTTCGCAATTAAG 1020  
 |||||||  
 Db 1127 GATGATTAATCAACCTTTGCGAGTTGATTAACGTCACCGTCACCTTCGCAATTAAG 1186  
 |||||||  
 QY 1021 CAACCTGTAATCTGCTGACGATCAAAATGGAATCTGCATGCATCCATTTGGCTTCGA 1080  
 |||||||  
 Db 1187 CAACCTGTAATCTGCTGACGATCAAAATGGAATCTGCATGCATCCATTTGGCTTCGA 1246  
 |||||||  
 QY 1081 ATGTTATGAGGTTTAAGCATTTGATTTGCTGCCGTCATTTATTTACCTTTGGTTTAA 1140  
 |||||||  
 Db 1247 ATGTTATGAGGTTTAAGCATTTGATTTGCTGCCGTCATTTATTTACCTTTGGTTTAA 1306  
 |||||||  
 QY 1141 TTTCGTCCAATCAGAAGCA 1161  
 |||||||  
 Db 1307 TTTCGTCCAATCAGAAGCA 1327  
 |||||||

## RESULT 8

; US-09-812-642-2/c  
 ; Sequence 2, Application US/09812642  
 ; Patent No. 6368600  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chandrasekar, Ramaswamy  
 ; APPLICANT: Morales, Tony H.  
 ; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
 ; FILE REFERENCE: HW-8  
 ; CURRENT APPLICATION NUMBER: US/09/812,642  
 ; CURRENT FILING DATE: 2001-03-20  
 ; PRIOR APPLICATION NUMBER: 09/323,427  
 ; PRIOR FILING DATE: 1999-06-01  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: Patentln Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1779  
 ; TYPE: DNA  
 ; ORGANISM: Dirofilaria immitis  
 ; US-09-812-642-2

Query Match 100.0%; Score 1161; DB 4; Length 1779;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Oy	1081	ANGTTATGGGTTTAAAGCATTCGATTGAATGCCTGCCGATCAATTAACATTTCCGTTTAA	1140
Dd	533	ATGTATTATGGTTTAAAGCATTCGATTGAATGCCTGCCGATCAATTAACATTTCCGTTTAA	474
Oy	1141	TTTCGTCCAATAACAGAGCA	1161
Dd	473	TTTCGTCCAATAACAGAGCA	453
RESULT 9			
US-09-323-427-6			
Sequence 6, Application US/09323427			
Patent No. 6248329			
GENERAL INFORMATION:			
APPLICANT: Chandrashekar, Ramaswamy			
APPLICANT: Morales, Tony H.			
TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid			
FILE REFERENCE: HW-8			
CURRENT APPLICATION NUMBER: US/09/323,427			
CURRENT FILING DATE: 1999-06-01			
EARLIER APPLICATION NUMBER: 60/087,435			
EARLIER FILING DATE: 1998-06-01			
NUMBER OF SEQ ID NOS: 18			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 6			
LENGTH: 1372			
TYPE: DNA			
ORGANISM: Dirofilaria immitis			
US-09-323-427-6			
Query Match                  43.4%; Score 504.2; DB 4; Length 1372;			
Best Local Similarity        76.9%; Pred. No. 3,7e-137;			
Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;			
Oy	1	ATGATGATTCGCTTATTCCTTTCGTACTACACTTATTCG-CATTCCTTATTCGATTC	59
Dd	32	ATTATCATTATTCCTCTTGTTATTCCTTATTCCTTATTCCTGACCTCGATTCGATTC	91
Oy	60	GCTTACACATGCTGTCGAAGGTGACCCGACAATTTGAATGTGACCAACTTCATAACAT	119
Dd	92	TATTACATATGCTGTCGAAGGTGACCTGTAATAGATGTGGCGACGCTTCGATTAACAT	151
Oy	120	CAATTTAATACAGCTAATGCTATTCGAGAAGCATGTTATGTGAAGAGCTTTATGATCA	179
Dd	152	CAATTTCAATACAGAAATACATTTGGAAGCACAGATATGAAAAGACCTGTATGATCA	211
Oy	180	AGAAGTTCGCTAATGATGAAGGTGACGTCAGTTCCGGAATTCCTTCATTTGCA	239
Dd	212	GGATCAATGCTGCTTCAGATAGTAAGGACGCGAGTAGCTGGAATTCGATTCGATGGA	271
Oy	240	TTTCATGATGTTGGCGCGATCGATCTGTGATCGACGTGATTTTTGTAAACAACAC	299
Dd	272	TTTCGTATATGTTGAACGATCAAGATCCTTAATCTCTGCTGTTTTGTAAACACTGT	331
Oy	300	TGTTTTCATTTGTTTTCATTCATTTATTTGTTACCAAAGTTGATCGTCGATTCGATCA	359
Dd	332	AGTTTCATTCATTTACATTCACAAATTCGTTTACAAAAAATATTCGACGATTCGATTA	391
Oy	360	ATGCTTTTACATG-G-AAGCTGATTAACAGTTAGTGC-ACAGATTGAGGTATCTGAATC	417
Dd	392	ATGTTTATATATGGAAGCTGATTAAGACCGTTAGTACTGTCCTTGAAGTATCTGAAATG	451
Oy	418	ACAATGCTTTTCAAACTCAATTTGCCGATCCGACAGTATGCCGTTATTCATAATTTTGGAT	477
Dd	452	ACTTACGATTCCTCAAACTCAAGTGTATCCAAAGCCGATATGTCGATATGAGATTTTGGAA	511
Oy	478	GGTGAACCAACCGGATCAACAGTTAAATTTGTTATCATTTGTGCACGAGTTATATCAATA	537
Dd	512	GGTGAACCAACCGGATCAACAGTTTGTGATTTGAAATGATTTGGAGATCATGATATATCAAAA	571
Oy	538	TGACATCGATTTTGAAAACCGTTGATACCTTTTGCGCGGTTGTCCATTCCTGCTTGTGC	597



Db	572	TGACATGTGATTCAGAGACATACAGATACATCTGTGCATTTAGTACATTCATGTGTTGTC	63.1
Qy	598	GATGATGCTAACGGTGTACTCTGTGGAATTTCTAAATGCTGATGATGTGCTCTTTGATAAA	65.7
Db	632	GATATGGAAAAAGGTGTAGTGCAGTGTGGAATTTCTGTAATAAAGAGATTTGCTTTGGACAAA	69.1
Qy	658	TATTGTCTAAATTAATTTGGAATATCCAAACAGATTTAATGGCTGGCCAAAGCTCAGCTA	71.7
Db	692	TATTTACTCATTAATTTGGAAATATATACAGATTTAATGGCTGGCCAAAGCTCATGTT	75.1
Qy	718	TACCAATATGCGGATTCGATCACAGCTTTTCTATCAATGCCAGATCAGTATTAACATTAAA	77.7
Db	752	TATATAATATGAGATCGATCCGAACCTTCTATCTCAATGCCAGATTTAGTATTAACAAATTTAAA	81.1
Qy	778	GAAACCAATATACGAATGTGTTGCGACCAATATGTCAGAACACACAGAGATTCGGAGCTGT	83.7
Db	812	GAGCCACATATACGAATGTCTCCGACCAACATATGCACAGAGCATCAGAGATTTGGTGCCTA	87.1
Qy	838	AAAAACAGG	84.5
Db	872	AAATCTGG	87.9

```

RESULT 10
US-09-323-427-7/c
Sequence 7, Application US/09323427
Patent No. 6248329
GENERAL INFORMATION:
APPLICANT: Chandrashekar, Ramaswamy
APPLICANT: Morales, Tony H.
TITLE OF INVENTION: Parasitic Helminth Cytolisin Proteins, Nucleic Acid
TITLE OF INVENTION: Molecules, and Uses Thereof
FILE REFERENCE: HW-8
CURRENT APPLICATION NUMBER: US/09/323.427
CURRENT FILING DATE: 1999-06-01
EARLIER APPLICATION NUMBER: 60/087,435
EARLIER FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 7
LENGTH: 1372
TYPE: DNA
ORGANISM: Dirofilaria immitis
US-09-323-427-7

```

Query Match	43.48;	Score 504.2;	DB 4;	Length 1372;
Best Local Similarity	76.98;	Pred. No. 3.7e-137;		
Matches 652; Conservative	0;	Mismatches 193;	Indels 3;	Gaps 3;

Qy	1	ATGTCGATTCGGTCTTAATGCGTTTCCTGACTACACTATATG-CATTGCTCTTAATGCATTCC	59
Db	1341	ATTATACATTAATTTCTCCCTGTGTAATCCCTACTCTTTTCACTGAGCCCTACGATACATGATCCC	1282
Qy	60	GGTTGACAAATGCTGTGGAAGGTGAGCCAGAAATGTAATGTGACCAACCTCAATTAACAAT	119
Db	1281	TATTTGACAAATGCTGTGGAAGGTGACCTGAAATAGATGTGGCGACGTTGATTAACAAT	1222
Qy	120	CAATTTTAAATACAGTAATGCAATTCGAAAGACATGTTATGTGAAGAGCTCTTATGATCA	179
Db	1221	CAATTTCAATACAGTAATGCAATTTGAAAGACAGCTATATGTGAAGAGCTCTATGATCA	1152
Qy	180	AGAAAGTTGCCGTAATGATGAAGGTGAGCGTCAAGTTGCCGGAATTTCACTTCATTTGGA	239
Db	1161	GGATGATATGTCGTTACGATAGTAATGAGACGGCAGGTGAGTAATGCAATTTGGCAATGGA	1102
Qy	240	TTTATGCAATGTTGCCCGTACAGCATCTCTGTAATCCAGCTGATATTTTGTAAACAAC	299
Db	1101	TTTCGTAAATGTTGAACGATACAGATCCCTTAATCTCGTGGTGTTTTGTAAACAACGT	1042
Qy	300	TGTTTCATTTTCGTTTCATCCATATATTTTTCACAAAGTTGATCGTGATATCGAGTACA	359
Db	1041	AGTTTGCATTTTACATTTTACCAAAATTCGTTTCAAAAATAGTACGAGTATTCGATACA	982

QY	360	ATGCTTTACATGC - AACGTATAAACAGTTAGTC - ACAGATTGAGGTATCTGAATC	417
Db	981	AGGTTTATATAGAAACCTGATTAAGACCGTTAGTCTGCTTGAAGTATCTGAAATG	922
QY	418	ACAACTGCTTTCAACATCAAAATTGTCGCCGATGCGAGTATGCCGTATGAAATTTGGAT	477
Db	921	ACTACAGCATTCGAAACCTCAAGTGGTACCATTGCCCGGATTTGCATATGAGATTTTGGAA	865
QY	478	GGTGAACCAACCCGTCAAACCACTTCATTTGCTCATTTGCTCAGCCAGTTTATCATATAA	533
Db	861	GGTGAACCAACCTGGTGCACCTGTTGATTTCCAATGATGGAGATTCATGTAATACACAAA	802
QY	538	TGGAGATCGAGTTTCGGAACCGTGTGATCTTTTGGCGGGTGTGCATTCCTGCTTTGTC	597
Db	801	TGGACATGTGATTGACGACACTACAGATATCATTTCTGTGATTAGTACATTCATGTGTTGTC	742
QY	538	GATGATGCTAAACGCTGATCTACTGTGGAAATTCATAATGCTGATGAGTGTGCTCTTGTATAAA	657
Db	741	GATGATGGAAGAAAGGTGATGAGCTAGTGGAGATCTCGATGAAAGAGATGTGCTTTGGACAAA	682
QY	658	TATTTGCTAAATTAATTGSGATATCCAAACATTAATTAAATGGCTGGCCAAAGAGCTACGTA	717
Db	681	TATTTACTCAATTAATTGGAATATATTTACAGATTTAAATGGCTGGCCAAAGAGCTACATGTT	622
QY	718	TACAAATATGGGATCGATTCACAGCTTTTCTATCAATGCCAGATCACTATTACCATATAA	777
Db	621	TATAATATATGACAGATCGATTCAGAACTTACTATCATGATGCCAGATATTGATATAACATATAA	562
QY	778	GAACCAATATACGAATGTGTTGCGACCAAAATGTTCAAAACCAAGATTCGGAGCTGTT	837
Db	561	GAGCCACATATGAGATGTCTCTGACCAACAAATGACACAGAGCCACAAAGATTTGGTCCATA	502
QY	838	AAAAACAG 845	
Db	501	AAATCTGG 494	

RESULT 11  
 US-09-812-642-6  
 ; Sequence 6, Application US/09812642  
 ; Patent No. 6368600  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chandrashekar, Ramaswamy  
 ; APPLICANT: Morales, Tony H.  
 ; TITLE OF INVENTION: Parasitic Helminth Cuticular Proteins, Nucleic Acid  
 ; FILE REFERENCE: Molecules, and Uses Thereof  
 ; FILE REFERENCE: HM-8  
 ; CURRENT APPLICATION NUMBER: US/09/812,642  
 ; CURRENT FILING DATE: 2001-03-20  
 ; PRIOR APPLICATION NUMBER: 09/323,427  
 ; PRIOR FILING DATE: 1999-06-01  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6  
 ; LENGTH: 1372  
 ; TYPE: DNA  
 ; ORGANISM: Dirofilaria immitis  
 ; US-09-812-642-6

Query Match	43.48;	Score 504.2;	DB 4;	Length 1372;
Best Local Similarity	76.98;	Pred. No. 3.7e-137;		
Matches 652; Conservative	0;	Mismatches 193;	Indels 3;	Gaps 3;

QY	1	ATGATGATTCGTCCTTATTCGCTTCCTGACCTACCTAATTCG-CATTGCTTAATTCGATTC	59
Db	32	ATTATCATTTATTCCTTCCTGTTATTCCTACTTCCTACCTGACCTACGATCATCGATCCC	91
QY	60	GCTTGACAAATCGTCGCGAAGTGGACCCGAATGATATGTCGACCACTTCAATACAAAT	119
Db	92	TATTGACATGCTGCGAAGTGGACCTGGAATATGAATGTGGCCACTTCGATTAACAAT	151
QY	120	CAATTTTAATACAGCTAATGATTGCAAGGACATGTTATGTGAAGGTCTTATATGATCA	179

Query Match	43.48;	Score 504.2;	DB 4;	Length 1372;
Best Local Similarity	76.98;	Pred. No. 3,7e-137;		
Matches 652;	Conservative 0;	Mismatches 193;	Indels 3;	Gaps 3;
1	ATGATGATTCGCTCTTATTCGCTTTCGTAATCTGACCTATTAATG-CATTTGCTATATTCGATTC	59		
Db	1341 ATTTACATTAATTCCTCTTCTGTTATTCCTACTTTTCTACTTGACCTACGATACATGATCCC	1282		
QY	60 GGTGTGACAAATGGTGTGCAAGGTGAGCCAGAAATGAAATGTTGGACCAACTTCCAAATTAACAT	119		
Db	1281 TATTGACAAATGGTGTGCAAGGTGAACTGAAATAGAAATGTGGCCAGCTTCGATTAACAT	1222		
QY	120 CAATTTATATTCACAGTAATGTCATTCGAAGCAGCATGTTATATGTGAAGGTCTTATGATCA	179		
Db	1221 CAATTTCAATTCCTGAAATATTAATCATTTTGAAAGCAGCATATATTTAAAGACCTATGATCA	1162		
QY	180 AGAAGTGTCCGCTGATGATGTAAGGTGAGCGACGATGTCGCGGAATTTGCATTCATTTGA	239		
Db	1161 GGATGAAATGTGCTTCAGATAGTAATGACGCGCAGGTAGCTGGAATTCGAATTTGGCAATGGA	1102		
QY	240 TTTACATGCAATGTTCCGGGTACAGCATCTGCAATTCAGCATCCAGTGTATTTTGTACACACAC	299		
Db	1101 TTTGCTGATATGTGACAGATCACGATCCCTTAATCTCCTGCTGTGTTTGTGAACACATGT	1042		
QY	300 TGTGTGCTATTTTCGTTTCAATTCATTTTGTTTACCAAAAGTATGTGTCATATGCAATGACA	359		
Db	1041 AGTTGTCTATTCATTTTCCTCCAAATTCGTTACCAAAATATGATATGACACATATCTATTAACA	982		
QY	360 ATGCTTTTACATGG-AAGCTGATATAACAGTTACTGTC-ACAGATTTGAGTATCGAATAC	417		
Db	981 ATGTTTATATATGGAAGCTGATAGACCGTTACTACTGCTCTGGAAGTATCGAATATG	922		
QY	418 ACAACTGCTTTTCAAACTCAAAATTTGCCGATGCGCAGTATGCCCTTATGAAATTTTGGAT	477		
Db	921 ACTACAGATTTCCAAACTCAAGTGTGTACCATGCGCGATGTGATATGAGATTTTGGAA	862		
QY	478 GGTGTGACCAACCGGTCAACACAGTTCAATTTGCTATTCATTTGTGACGCCAGTTATCATATA	537		
Db	861 GGTGTGACCAACCTGTGTGTCACCTGTTCGATTTGCAATGATTTGAGATATCATATACACAAA	802		
QY	538 TTGACATGCGCGATTTGGAACCGTTGATATCTTGCGCGGGTGTGCAATTCCTCGCTTGTGC	597		
Db	801 TTGACATGCTGTTAGAGACTTACAGATATCATTTCTGTGCTATTTAGTACATTCATGTGCTTGTG	742		
QY	598 GATGATGGTAAACGGTGTACTGTTGAAATTTCTAATCTGATGATGTGCTCTTGTATATA	657		
Db	741 GATGATGGAAAGAGGATAGTACAGTGTGAGATTTCTGAAATGAAGAAGATGTGCTTTGGACAAA	682		
QY	658 TATTTGCTAAATTAATTTTGGAAATATCCAACAGATTTAATGCTGTGGCCAGAAAGCTCACGTA	717		
Db	681 TATTTACTCAATTAATTTTGGAAATATATTAACAAATTTAATGGTGGCCAGAAAGCTCAATGTT	622		
QY	718 TACAAATATGTGGATGATCATCACAGCTTTTGTATCAATGGCCAGATCAATTAACATTAATA	777		
Db	621 TATTAATATGTGATGATCATCATCGAAGCTTTACTATTAATGGCCAGATTAATATTAACATTAATA	562		
QY	778 GAACCAATATGACGAATGTTTGAACACCAATGTTTCAGAACCCACAAGGATTTGCGAGCTGT	837		
Db	561 GAGCCACATATACGAATGTCTCGACCAACAATGACAGAGCCACAAGGATTTGTTGGCCATA	502		
QY	838 AAAACAGG 845			
Db	501 AAATCTGG 494			
RESULT 13				
US-09-323-427-16				
; Sequence 16; Application US/09323427				
; Patent No. 6248329				
; GENERAL INFORMATION:				
; APPLICANT: Chandreshkar, Ramaswamy				
; APPLICANT: Morales, Tony H.				









Db 200 GCACGGTGTATATACCGAAGAGATCCGAACTTCTCAGCAAT 157

## RESULT 21

US-09-308-003-45  
; Sequence 45, Application US/09308003  
; Patent No. 6326170  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin K. R.  
; APPLICANT: Lonetto, Michael A.  
; APPLICANT: Warren, Patrick V.  
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,  
; TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES  
; FILE REFERENCE: GM10093  
; CURRENT APPLICATION NUMBER: US/09/308,003  
; CURRENT FILING DATE: 1999-05-10  
; EARLIER APPLICATION NUMBER: 60/058,710  
; EARLIER FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 348  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(348)  
US-09-308-003-45

## Query Match

Best Local Similarity 3.0%; Score 35.4; DB 4; Length 348;  
Best Local Similarity 53.2%; Pred. No. 0.6;  
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 16 ATTCCTTCTGTACTACTATTCATGATGCTTATTCGATCCGGTGCACATGCTGC 75

Db 82 ATTCCTTACCGTATGAAGAGTATGACGTTTATTTATTAATCCGATGCAATGCGTGA 141

Qy 76 GAAGTGAGCCAGAAATGATGTGGACCACTTCAATTAACATGATTTTAATACAGT 135

Db 142 AAGGTATGCGCTGAAGAGATATGATTTGGTGGACAAAGTAACGTGCAATCGCAA 201

Qy 136 AATGCATTGGAAGACATGTT 156

Db 202 CAAGCAGCTCTTGCAAAATGCT 222

## RESULT 22

US-09-308-003-8  
; Sequence 8, Application US/09308003  
; Patent No. 6326170  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin K. R.  
; APPLICANT: Lonetto, Michael A.  
; APPLICANT: Warren, Patrick V.  
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,  
; TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES  
; FILE REFERENCE: GM10093  
; CURRENT APPLICATION NUMBER: US/09/308,003  
; CURRENT FILING DATE: 1999-05-10  
; EARLIER APPLICATION NUMBER: 60/058,710  
; EARLIER FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 1182  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1179)  
US-09-308-003-8

Query Match 3.0%; Score 35.4; DB 4; Length 1182;  
Best Local Similarity 49.2%; Pred. No. 1.1;  
Matches 90; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 16 ATTCCTTCTGTACTACTATTCATGATGCTTATTCGATCCGGTGCACATGCTGC 75

Db 805 ATTCCTTACCGTATGAAGAGTATGACGTTTATTTATTAATCCGATGCAATGCGTGA 864

Qy 76 GAAGTGAGCCAGAAATGATGTGGACCACTTCAATTAACATGATTTTAATACAGT 135

Db 865 AAGGTATGCGCTGAAGAGATATGATTTGGTGGACAAAGTAACGTGCAATCGCAA 924

Qy 136 AATGCATTGGAAGACATGTT 156

Db 925 CAAGCAGCTCTTGCAAAATGCT 945

## RESULT 23

US-08-894-017-24  
; Sequence 24, Application US/08894017  
; Patent No. 6024958  
; GENERAL INFORMATION:  
; APPLICANT: Lehner, Thomas  
; APPLICANT: Kelly, Charles  
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE  
; TITLE OF INVENTION: OF COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1888  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,017  
; FILING DATE: 31-JUL-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB96/00207  
; FILING DATE: 31-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 25150-20067.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-887-1500  
; TELEFAX: 202-822-0168  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4865 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
US-08-894-017-24

Query Match 2.9%; Score 34.2; DB 3; Length 4865;

Best Local Similarity 49.2%; Pred. No. 5.2;  
Matches 90; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 601 GATGTAACGGTATGATGTAATTCATGATGATGATGCTCTGATTAATAT 660

Db 1426 GATGTAACGGTATGATGTAATTCATGATGATGATGCTCTGATTAATAT 1485

Qy 661 TTGCTAATATTTGGAATATCCACAGATTTAATGCTGCCAAGAGCTCAGCTATAC 720



Db 1486 CAAAAGATTAGACAGACTATCCAGTTAAGTTAAAGCATACGAAGATGAACAGCTTCT 1545  
QY 721 AATATGCGATGCATACACGCTTTCTATCAATGCCAGATCAGATATACATTAAGAA 780  
Db 1546 ATTAAGCTGCATGCGAAGACTTGAAACATTAATAATGAAGACGAAACTTAACAGAA 1605  
QY 781 CCA 783  
Db 1606 CCA 1608

RESULT 24  
US-09-323-427-11  
; Sequence 11, Application US/09323427  
; Patent No. 6248329  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HW-8  
; CURRENT APPLICATION NUMBER: US/09/323,427  
; CURRENT FILING DATE: 1999-06-01  
; EARLIER APPLICATION NUMBER: 60/087,435  
; EARLIER FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 34  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-323-427-11

Query Match 2.9%; Score 34; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 729  
Db 1 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 34

RESULT 25  
US-09-323-427-12/C  
; Sequence 12, Application US/09323427  
; Patent No. 6248329  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HW-8  
; CURRENT APPLICATION NUMBER: US/09/323,427  
; CURRENT FILING DATE: 1999-06-01  
; EARLIER APPLICATION NUMBER: 60/087,435  
; EARLIER FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 34  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-323-427-12

Query Match 2.9%; Score 34; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 729  
Db 34 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 1

RESULT 26  
US-09-812-642-11  
; Sequence 11, Application US/09812642  
; Patent No. 6368600  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HW-8  
; CURRENT APPLICATION NUMBER: US/09/812,642  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 09/323,427  
; PRIOR FILING DATE: 1999-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 34  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-812-642-11

Query Match 2.9%; Score 34; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 729  
Db 1 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 34

RESULT 27  
US-09-812-642-12/C  
; Sequence 12, Application US/09812642  
; Patent No. 6368600  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HW-8  
; CURRENT APPLICATION NUMBER: US/09/812,642  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 09/323,427  
; PRIOR FILING DATE: 1999-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 34  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-812-642-12

Query Match 2.9%; Score 34; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 729  
Db 34 GGCTGGCCAGAGCTCAGTATCAAAATATGCG 1



```

1  REGISTRATION NUMBER: 29,768
2  REFERENCE/DOCKET NUMBER: 16754/115 CHCO
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: (703)836-9300
5  TELEFAX: (703)683-4109
6  TELEEX: 899149
7  INFORMATION FOR SEQ. ID NO: 1:
8  SEQUENCE CHARACTERISTICS:
9  LENGTH: 1867 base pairs
10  TYPE: NUCLEIC ACID
11  STRANDEDNESS: double
12  TOPOLOGY: linear
13  MOLECULE TYPE: DNA (genomic)
14  ORIGINAL SOURCE:
15  STRAIN: 164A-1
16  FEATURE:
17  NAME/KEY: CDS
18  LOCATION: 716..1849
19  OTHER INFORMATION: /product="mature protein of
20  OTHER INFORMATION: 164A-1"
21  FEATURE:
22  NAME/KEY: sig_peptide
23  LOCATION: 716..1021
24  FEATURE:
25  NAME/KEY: mat_peptide
26  LOCATION: 1022..1846
27  OS-07-772-087-1

```

```
US-09-385-982-45
; Sequence 45, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(630)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-45

Query Match          2.8%; Score 32.6; DB 4; Length 630;
Best Local Similarity 48.6%; Pred. No. 5.3;
Matches 89; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 614 ATACTGTGGAATTTCTAAATGCTGATGATGTGCTTGTGATAAATTTGCTAAATAT 673
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 94 ATACTGTTTCATTTGATTCAGAGGGCTTGTGTAAGAAATAGTTCAAAATCATT 153

QY 674 TGGAAATATCCACAGATTTAATGGCGGCCAGAAAGCTACGATATCAATATGCGGATC 733
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154 GAAACCCAGAACAGAAATGCTTATGTTATCTGTGATACACTGTTTTCTAAACCAAG 213

QY 734 GATCACAGCTTTCTATCATGCGCAGATCATTTACATTAAGAACCAATAGCGAAT 793
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 214 GATTTTCTTTTATTAATATGCAACATAGACATGCCATTAACAGAAATTAACCAAT 273

QY 794 GTG 796
    |||
Db 274 GTG 276

RESULT 35
5352450-1
; Patent No. 5352450
; APPLICANT: KOGA, TOSHIHIKO; OKAHASHI, NOBUO; TAKAHASHI, ICHIRO;
; SHIUYA, KOJI; OHYA, HIROTAKA
; TITLE OF INVENTION: METHOD FOR PREPARING VACCINE FOR DENTAL
; CARIES AND VACCINAL COMPOSITIONS FOR DENTAL CARIES USED AS NASAL
; DROP
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/529,602
; FILING DATE: 29-MAY-1990
; SEQ ID NO: 1
; LENGTH: 5177
5352450-1

Query Match          2.8%; Score 32.6; DB 6; Length 5177;
Best Local Similarity 48.6%; Pred. No. 16;
Matches 89; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
```

```
QY 661 TGCCTAAATATTTGGAAATATCCACAGATTTAATGCTGGCCAGAGCTACGATATAC 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1586 CAAAAGATTTCACAGATCATCGATTAAAGTAAAGCATACGAAATGAACAACCTTCT 1645

QY 721 AATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1646 ATTAAGCTGCTACTGCGAGAACTTGAATAAATGAAGAGCGAAACTTAACAGAA 1705

QY 781 CCA 783
    |||
Db 1706 CCA 1708

RESULT 36
PCT-US96-05320A-705/c
; Sequence 705, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20003-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April 22, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steife
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.014PC01
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 705:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US96-05320A-705

Query Match          2.8%; Score 32.4; DB 5; Length 576;
Best Local Similarity 56.6%; Pred. No. 5.8;
Matches 60; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
QY 1038 TGCAGTACAAATGGAATGCAATGTCAGTACCAATTGGCTTCAATGTTATGAGTTAAG 1097
```



```
SEQUENCE CHARACTERISTICS:
LENGTH: 4090 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 197..1699
OTHER INFORMATION: /note="R011 sequence of longest
OTHER INFORMATION: open reading frame; other possible start codons ATG/met4;
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 197..1699
US-08-694-078-5

Query Match      2.8%; Score 32.4; DB 4; Length 4090;
Best Local Similarity 44.5%; Pred. No. 16;
Matches 129; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 150 ACATGTTATGTGAAGAGTCTTTATGATCAAGAGGTTCCCGTAATGATGAAGGTGGAGC 209
DB 2895 ACATGTTTGAATAACGTCACATGTAAGAAATACAGACAGATGATATCTCATTA 2836
QY 210 TCAGATGCCGGAATTTCACTTCATTCATTCAGCAATGTTGCGCGTACAGATCTCT 269
DB 2835 ACAAGTTAATCTGTGAGAGACACATTTGACTCTCAGACAGATGAATAATGTGGGAA 2776
QY 270 GAATCCAGCTGATTTTGTAAACAACACTGTTCATTTGCTTACATTCATTTATTTCT 329
DB 2775 ATATCGATGCGCAACCCCTTGTAATGAAGGGTTTCGGCATTTTATCTTATTAAT 2716
QY 330 TACCAAGTTGATCGTCATATGATACATGCTTTTACATGAGAGCGATAAACAGT 389
DB 2715 TACAAGAAGATGTTTGTGAACCCGTGAGAGGCAATTAAGTGTGATCAAGAGGAGG 2656
QY 390 TAGTGCACGATTTGAGATATCTGAANAACAACTGCTTTTCAACTCAAA 439
DB 2655 AGAAGAAACCCGCGGTGTTTCGCGCCACAATATCTTGTGAATACAGA 2606

RESULT 39
US-09-453-702B-11/C
; Sequence 11, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plunkney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296, 95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31960
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-453-702B-11

Query Match      2.8%; Score 32.4; DB 4; Length 31960;
Best Local Similarity 52.9%; Pred. No. 45;
Matches 91; Conservative 1; Mismatches 77; Indels 3; Gaps 1;

QY 1 ATGATGATTCGCTTATTCCTTCTGCTACTATGCACTATTCATTCGATTCG 60
DB 1051 ATATTTACTGCTGACATTCATGACAAAGGCTGCGATTCGATTCATGCT 992
QY 61 GTTGACATGCTGTCGAAAGTGCAGAA--ATTGAATGACCACTTCAATACA 117
DB 991 AAGAATGATGAGTGTTGATGACACTTGAATTTATGATTTGTTAAACACTTAT 932
QY 118 ATCAATTTAATACAGTAATGATTCATTCGAAGCATGTTTATGTGAAGGTC 169
DB 931 ATGAGCACTATTACACTTAAAGCAAGCAAAATATTAAGAGAAATTC 880

RESULT 40
US-09-134-001C-1411
; Sequence 1411, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1411
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1411

Query Match      2.8%; Score 32.2; DB 4; Length 1860;
Best Local Similarity 45.0%; Pred. No. 12;
Matches 121; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 784 AATAGCGAATGTTGACACATGTTCAAGCAAGGATTTGGAGCTGTTAAACA 843
DB 88 AATGCCAAGATTAATCTAATTAAGAGTTAGAAAATTAAGATATCGTGTATGACTCA 147
QY 844 GGTGTGCCGAGCAAAACCTGCTGACCTGCGCACTTCGTTTACTCAAGAAAAGATCT 903
DB 148 GCAGGTATGCTGATGTAATGATGATGTACAAAACCTTTTAAAGAAAAGGAATT 207
QY 904 CGAAGACGAGAAATATCATTTGATGATGATATCAACACCCCTGGAATGCGAT 963
DB 208 GCTGAATTAAGTAAAGTTCGAGATTAATGATGATGAGATGTTAGCTTATGAGATTTGCTAT 267
QY 964 GATAATCAAGCTTCCAGTTGATTTACGTACCGGTGACCTTGCACAATATGACAA 1023
```

Db 268 ACAAGTGGGCGACACATGCGTCCAAATTATGAAATTCACATCCACACCGAGTCACA 327  
 Oy 1024 CCTGTAACTCTGCTGCAGTACAAAATG 1052  
 Db 328 TCTGACGTTTACATTAAGTTCATAATG 356

Search completed: March 28, 2003, 17:00:55  
 Job time : 141 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 16:45:05 ; Search time 120 Seconds

(without alignments)  
8233.163 Million cell updates/sec

Title: US-10-054-562a-3

Perfect score: 1161  
Sequence: 1 atgatgattcgtctatttcgc.....tttcgcaaatcagaagca 1161

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161	100.0	1161	9	US-10-054-562a-3
2	1161	100.0	1161	9	US-10-054-562a-5
3	1161	100.0	1179	9	US-10-054-562a-1
4	1161	100.0	1179	9	US-10-054-562a-2
5	504.2	43.4	1372	9	US-10-054-562a-7
6	504.2	43.4	1372	9	US-10-054-562a-16
7	461	39.7	892	9	US-10-054-562a-18
8	461	39.7	892	9	US-10-054-562a-18
9	310.4	26.7	813	9	US-10-054-562a-8
10	310.4	26.7	813	9	US-10-054-562a-10
11	40.2	3.5	2000	9	US-09-938-842a-3231
12	38	3.3	6799	9	US-09-902-941-1883
13	38	3.3	6799	9	US-09-849-626-1883
14	38	3.3	6799	9	US-10-017-754-1883
15	36.2	3.1	510	10	US-09-864-761-18737
16	36.2	3.1	1103	10	US-09-974-300-5151
17	35.4	3.0	7563	7	US-08-781-986a-33
18	35.4	3.0	378361	9	US-09-901-136-3
19	35.2	3.0	465	10	US-09-864-761-1160

20	35	3.0	660	10	US-09-864-761-19488	Sequence 19488, A
21	35	3.0	962	10	US-09-864-761-2772	Sequence 2772, A
22	34.8	3.0	583	10	US-09-864-761-20772	Sequence 20772, A
23	34.8	3.0	1608	10	US-09-815-242-3929	Sequence 3929, Ap
24	34.8	3.0	1611	10	US-09-815-242-6630	Sequence 6630, Ap
25	34.8	3.0	1959	10	US-09-864-761-4012	Sequence 4012, Ap
26	34.8	3.0	6617	9	US-10-210-296-1	Sequence 1, Appl1
27	34.8	3.0	8395	10	US-09-070-927a-217	Sequence 217, App
28	34.6	3.0	1479	10	US-09-070-927a-405	Sequence 405, App
29	34.6	3.0	21252	10	US-09-070-927a-94	Sequence 94, Appl
30	34.2	2.9	402	10	US-09-960-352-11958	Sequence 11958, A
31	34.2	2.9	3549	10	US-09-070-927a-859	Sequence 859, App
32	34	2.9	34	9	US-10-054-562a-11	Sequence 11, Appl
33	34	2.9	34	9	US-10-054-562a-12	Sequence 12, Appl
34	34	2.9	325	10	US-09-864-761-25339	Sequence 25339, A
35	34	2.9	556	10	US-09-864-761-8619	Sequence 8619, Ap
36	34	2.9	895	10	US-09-745-288-91	Sequence 91, Appl
37	34	2.9	2322	9	US-09-764-872-725	Sequence 725, App
38	34	2.9	6035	7	US-08-781-986a-272	Sequence 272, App
39	33.8	2.9	1170	10	US-09-815-242-4340	Sequence 4340, Ap
40	33.8	2.9	1182	10	US-09-815-242-8304	Sequence 8304, Ap
41	33.8	2.9	32768	10	US-09-070-927a-100	Sequence 100, App
42	33.6	2.9	2000	9	US-09-938-842a-4475	Sequence 4475, Ap
43	33.6	2.9	4239	7	US-08-781-986a-333	Sequence 333, App
44	33.4	2.9	3327	10	US-09-070-927a-302	Sequence 302, App
45	33.4	2.9	16950	9	US-10-114-170-166	Sequence 166, App

#### ALIGNMENTS

RESULT 1  
US-10-054-562a-3  
Sequence 3, Application US/10054562a  
Patent No. US20020165375a1  
GENERAL INFORMATION:  
APPLICANT: Chandrasekar, Ramaswamy  
APPLICANT: Morales, Tony H.  
TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES,  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: HW-8-2  
CURRENT APPLICATION NUMBER: US/10/054, 562a  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: 09/812,642  
PRIOR FILING DATE: 2001-03-20  
PRIOR APPLICATION NUMBER: 09/323,427  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 1161  
TYPE: DNA  
ORGANISM: Dirofilaria immitis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1161)  
OTHER INFORMATION:  
US-10-054-562a-3  
Query Match 100.0%; Score 1161; DB 9; Length 1161;  
Best Local Similarity 100.0%; Pred. No. 4,7e-294;  
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ATGATGATTCGCTTATTCGCTTCTGAGACACTTATTCGATTCGATTCG 60  
DB 1 ATGATGATTCGCTTATTCGCTTCTGAGACACTTATTCGATTCGATTCG 60  
OY 61 GTTGACATGCTGTCGAGAGGAGCCAGAAATGAAATGAGCAATTCATCAATC 120  
DB 61 GTTGACATGCTGTCGAGAGGAGCCAGAAATGAAATGAGCAATTCATCAATC 120  
OY 121 AATTTTATACACGATTCGATTCGAGAGCAATTCGATTCGAGAGGCTTTATGACAA 180  
DB 121 AATTTTATACACGATTCGATTCGAGAGCAATTCGATTCGAGAGGCTTTATGACAA 180



QY	721	AAATATGCGATGATCAACAAGCTTTCTTCAATCAATGCGAGATACGATTAATACATTAAAGAA	780
Db	441	AAATATGCGGATGATCAACGCTTTTCTATCAATGCGAGATACGATTAATTAAGAA	382
QY	781	CCAAATGCGAATGTGTGACACCAATATTCAGAACCAAGATTTGGAGCTGTAA	840
Db	381	CCAAATGCGAATGTGTGACCAACAATGTTCAGAACCAAGATTTGGAGCTTTAA	322
QY	841	ACAGGTGTCGCCGACGAAAACCTGCTGCACGCTGGCAACTTCGTTTACTCAAGAAAAGA	900
Db	321	ACAGGTGTCGCCGACGAAAACCTGCTGCACGCTGGCAACTTCGTTTACTCAAGAAAAGA	262
QY	901	TCTGCAACACCGGAGAAATATCATTTGATACGAACTGATTAACAACCTTGAATTAAGC	960
Db	261	TCTGCAACACCGGAGAAATATCATTTGATACGAACTGATTAACAACCTTGAATTAAGC	202
QY	961	GATGATATATCAAGCTTTGCCAGTTGATTAACGTACACCGTGCACTTCTGCAACATTAATGA	1020
Db	201	GATGATATATCAAGCTTTGCCAGTTGATTAACGTACACCGTGCACTTCTGCAACATTAATGA	142
QY	1021	CAACCTTATACCTGCTGAGTACAAATGGAATGTCATGTCACCAATTTGGCTTCGA	1080
Db	141	CAACCTTATACCTGCTGAGTACAAATGGAATGTCATGTCACCAATTTGGCTTCGA	82
QY	1081	ATGTTTATGAGGTTTAAAGCATTTGATGATTTGCTGCCGTATTAATACATTTTCGTTTAA	1140
Db	81	ATGTTTATGAGGTTTAAAGCATTTGATGATTTGCTGCCGTATTAATACATTTTCGTTTAA	22
QY	1141	TTTTCGTCCAATCAGAAGGCA	1161
Db	21	TTTTCGTCCAATCAGAAGGCA	1

```

RESULT 3
US-10-054-562A-1
: Sequence 1, Application US/10054562A
: Patent No. US20020165375A1
: GENERAL INFORMATION:
: APPLICANT: Chandrashekar, Ramaswamy
: APPLICANT: Morales, Tony H.
: TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES, AND
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: HW-8-2
: CURRENT APPLICATION NUMBER: US/10/054,562A
: CURRENT FILING DATE: 2002-06-10
: PRIOR APPLICATION NUMBER: 09/812,642
: PRIOR FILING DATE: 2001-03-20
: PRIOR APPLICATION NUMBER: 09/323,427
: PRIOR FILING DATE: 1999-06-01
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1779
: TYPE: DNA
: ORGANISM: Dirofilaria immitis
US-10-054-562A-1

```

	Query Match	100.0%	Score 1161	DB 9	Length 1779	
	Best Local Similarity	100.0%	Pred. No. 6e-294			
	Matches 1161	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1 ATGATGATTCGCTCTATTGCTTCTTGTACTACACTTATGCACTTGCATTCGATTCG 60					
Db	167 ATGAGATTCCTGCTATTGCTTCTTGTACTACACTTATGCACTTGTGCTATTGCAATTCG 226					
Qy	61 GTTGCAATGCTGTGCGAAGGTGAGCCAGAAATGGAATGTGACCAACTTCAATTAACAATC 120					
Db	227 GTTGACAAATGGGTGCGAAGGTGAGCCAGAAATGGAATGTGACCAACTTCAATTAACAATC 286					
Qy	121 AATTTTAATACACGTAATATGCAATTCGAAAGACATTTTATATGTAAGGCTTTATATGTCAA 180					
Db	287 AATTTTAATACACGTAATATGCAATTCGAAAGACATTTTATATGTAAGGCTTTATATGTCAA 346					

OY	181	GAGGTTTCCCGTAATGATGAAGGTGGACGTCAAGTTCGCCGAATTTCACTCCATTGGAT	240
Db	347	GAAAGTTTCCCGTAATGATGAAGGTGGACGTCAAGTTCGCCGAATTTCACTCCATTGGAT	406
OY	241	TCATCAGTATGTTGGCCGCTACACAGATCTGTGAATCCACCGTGTATTTTGGTAACAACAAC	300
Db	407	TCATCAGTATGTTGGCCGCTACACAGATCTGTGAATCCACCGTGTATTTTGGTAACAACAAC	466
OY	301	GTTGTCAATTCGTTTCATTCATATTTGTGTACCAAAAGTTGATGCGTCATATCGAGTACAA	360
Db	467	GTTGTCAATTCGTTTCATTCATATTTGTGTACCAAAAGTTGATGCGTCATATCGAGTACAA	526
OY	361	TGCTTTTTCATGGAAGGTGATAAAAACAATTTAGTGTGACAGATGTTGAGTATTCGAANAACA	420
Db	527	TGCTTTTTCATGGAAGGTGATAAAAACAATTTAGTGTGACAGATGTTGAGTATTCGAANAACA	586
OY	421	ACTGCTTTTCAAACTCAAAATTTGTCGCGATGTCAGATGTCGCTTATGAAATTTTGGATGGT	480
Db	587	ACTGCTTTTCAAACTCAAAATTTGTCGCGATGTCAGATGTCGCTTATGAAATTTTGGATGGT	646
OY	481	GGACCAACCCGTCACACCGCTTCATTTCTATCATTTGTCACGCCAGTTTATCATTAATCG	540
Db	647	GGACCAACCCGTCACACCGCTTCATTTCTATCATTTGTCACGCCAGTTTATCATTAATCG	706
OY	541	ACATCGATTCGGAACCGGTGATATCTTTGCGCGGGTTCGCAATTCCTGCTTTGTGCGAT	600
Db	707	ACATCGATTCGGAACCGGTGATATCTTTGCGCGGGTTCGCAATTCCTGCTTTGTGCGAT	766
OY	601	GATGTTAACCGGTGATATCTGTGGAAATTTCTAAATGCTGATGATGTGCTTTGATTAATAT	660
Db	767	GATGTTAACCGGTGATATCTGTGGAAATTTCTAAATGCTGATGATGTGCTTTGATTAATAT	826
OY	661	TTTGCTAAATTAATTTGGATATTCACAACAGATTTAAATGGTGGCCAGAAGCTCAGGTATAC	720
Db	827	TTTGCTAAATTAATTTGGATATTCACAACAGATTTAAATGGTGGCCAGAAGCTCAGGTATAC	886
OY	721	AAATATGGGAGTCGATCACTACCTTTTCTATCAATATCCAGATCAGTATTAACCTTAAGAA	780
Db	887	AAATATGGGAGTCGATCACTACCTTTTCTATCAATATCCAGATCAGTATTAACCTTAAGAA	946
OY	781	CCAAATAGCGAATGTGTTTCGACACACATGTTTCAGAACCCAGAGATTTGGAAGCTGTTAAA	840
Db	947	CCAAATAGCGAATGTGTTTCGACACACATGTTTCAGAACCCAGAGATTTGGAAGCTGTTAAA	1006
OY	841	ACAGTGTGGTGGCGGACGCAAAACCTGCTCAGCTGGCGCAACTTCGTTTACTCAAGAAAAA	900
Db	1007	ACAGTGTGGTGGCGGACGCAAAACCTGCTCAGCTGGCGCAACTTCGTTTACTCAAGAAAAA	1066
OY	901	TCTGCAGAACCGGAGAAATATCATTTGATGTAGAACTGATATCAACACCCCTTGAATTAGC	960
Db	1067	TCTGCAGAACCGGAGAAATATCATTTGATGTAGAACTGATATCAACACCCCTTGAATTAGC	1126
OY	961	GATGATTAATCAAGCTTTTGCCAGGTGATTTAGCTACCGGTGCACCTTCTGCAACATTAATGGA	1020
Db	1127	GATGATTAATCAAGCTTTTGCCAGGTGATTTAGCTACCGGTGCACCTTCTGCAACATTAATGGA	1186
OY	1021	CAACCTGTAATACCTGTCTGCAGTACAAATATGAAATCTGTCAGTCAACACTTGTGGCTTCCA	1080
Db	1187	CAACCTGTAATACCTGTCTGCAGTACAAATATGAAATCTGTCAGTCAACACTTGTGGCTTCCA	1246
OY	1081	ATGTTTATGGGTTTAAGCATTTGCAATGATGATGTCTGCGCTCATTTATTACCAATTTGCTTTAAA	1140
Db	1247	ATGTTTATGGGTTTAAGCATTTGCAATGATGATGTCTGCGCTCATTTATTACCAATTTGCTTTAAA	1306
OY	1141	TTTGGTCCAAATCAGAAAGCA 1161	
Db	1307	TTTGGTCCAAATCAGAAAGCA 1327	

RESULT 4  
US-10-054-562A-2/c  
; Sequence 2, Application US/10054562A

```

; Patent No. US20020165375A1
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: PARASTIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: HW-8-2
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Dirofilaria immitis
US-10-054-562a-2

```

Query Match 100.0%; Score 1161; DB 9; Length 1779;

Best Local Similarity 100.0%; Pred. No. 6e-294;

Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATGATTCGCTTATTTGCTTTCTGCTACTACACTTATTCGATTTGCTTATTCGATTCG 60
DB 1613 ATGATGATTCGCTTATTTGCTTTCTGCTACTACACTTATTCGATTTGCTTATTCGATTCG 1554
OY 61 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGATGTGGACCACTTCAATACAAATC 120
DB 1553 GTTGACAATGGTGTGCAAGGTGAGCCAGAAATGATGTGGACCACTTCAATACAAATC 1494
OY 121 AATTTTATACAGTATGATCGAAGACATGTTATGTGAAGGCTTTATGATCA 180
DB 1493 AATTTTATACAGTATGATCGAAGACATGTTATGTGAAGGCTTTATGATCA 1434
OY 181 GAAGGTGCGGTATGATGAAGGTGAGCTCAAGTTGCCGGAATTTCACTTCATTTGAT 240
DB 1433 GAAGGTGCGGTATGATGAAGGTGAGCTCAAGTTGCCGGAATTTCACTTCATTTGAT 1374
OY 241 TCATGCAATGTCGGCGTACAGATCTCGAATCCACGCGGTATTTTGTAAACAACACT 300
DB 1373 TCATGCAATGTCGGCGTACAGATCTCGAATCCACGCGGTATTTTGTAAACAACACT 1314
OY 301 GTTGATATTCGTTTCATCATTTATTTTACCAAGTTGATCGTATGAGTACAA 360
DB 1313 GTTGATATTCGTTTCATCATTTATTTTACCAAGTTGATCGTATGAGTACAA 1254
OY 361 TCGTTTACATGGAAGCTGATAAACAGTTAGTGACAGATTGAGGTATCTGAAATCACA 420
DB 1253 TCGTTTACATGGAAGCTGATAAACAGTTAGTGACAGATTGAGGTATCTGAAATCACA 1194
OY 421 ACTGCTTTCAACATCAATTTGCCGATGCCAGATGCGGTATGCAATTTTGGATGCT 480
DB 1193 ACTGCTTTCAACATCAATTTGCCGATGCCAGATGCGGTATGCAATTTTGGATGCT 1134
OY 481 GGACCAACCGGTCAACCACTTATGCTATCATTTGTCAGCCAGTTATATCAATAATGG 540
DB 1133 GGACCAACCGGTCAACCACTTATGCTATCATTTGTCAGCCAGTTATATCAATAATGG 1074
OY 541 ACATGCGATTCGAAACCGTTGATCTTTCTGCGGGTGTGCCATTCCTGCTTGTGCTAT 600
DB 1073 ACATGCGATTCGAAACCGTTGATCTTTCTGCGGGTGTGCCATTCCTGCTTGTGCTAT 1014
OY 601 GATGTAACGGTGATACGTGTGAATTTCTAATGCTGATGATGTGCTTCTGATTAATAT 660
DB 1013 GATGTAACGGTGATACGTGTGAATTTCTAATGCTGATGATGTGCTTCTGATTAATAT 954
OY 661 TTGCTAAATTAATTTGGAATATCCACAGATTAATGGCGGCAAGAAAGCTACGTATAC 720
DB 953 TTGCTAAATTAATTTGGAATATCCACAGATTAATGGCGGCAAGAAAGCTACGTATAC 894

```

```

OY 721 AATATGCGGATGATCACAGCTTTTCTATCAATGCCAGATCAGTATACATTAAGAA 780
DB 893 AATATGCGGATGATCACAGCTTTTCTATCAATGCCAGATCAGTATACATTAAGAA 834
OY 781 CCAATAGGGAATGTGTGTGACCAATGTTCGAACCCACAAAGATTCGAGCTGTTAA 840
DB 833 CCAATAGGGAATGTGTGTGACCAATGTTCGAACCCACAAAGATTCGAGCTGTTAA 774
OY 841 ACAGGTGTGCGCCAGCAAAACCTGCTGACGGCGCACTTGGTTACTCAAGAAAGA 900
DB 773 ACAGGTGTGCGCCAGCAAAACCTGCTGACGGCGCACTTGGTTACTCAAGAAAGA 714
OY 901 TCTGCAAAACCGGAGAAATATCATTTGATGACGAATGATATCAACCCCTTGAATAC 960
DB 713 TCTGCAAAACCGGAGAAATATCATTTGATGACGAATGATATCAACCCCTTGAATAC 654
OY 961 GATGATTAATCAAGCTTTGCGAGTTGATTTACGTACACCGGTGCACTTCTGCAATATGA 1020
DB 653 GATGATTAATCAAGCTTTGCGAGTTGATTTACGTACACCGGTGCACTTCTGCAATATGA 594
OY 1021 CAACCTGTAATACCTTCTGACAGTACAAATGGAATGCAATGCACTTGGCTTCTCA 1080
DB 593 CAACCTGTAATACCTTCTGACAGTACAAATGGAATGCAATGCACTTGGCTTCTCA 534
OY 1081 ATGTTATGCGTTTAAAGATTCGATGATTTGCTGCCGCTATTTACATTTGCTTTAA 1140
DB 533 ATGTTATGCGTTTAAAGATTCGATGATTTGCTGCCGCTATTTACATTTGCTTTAA 474
OY 1141 TTTTCGCAAAATTCAGAAAGCA 1161
DB 473 TTTTCGCAAAATTCAGAAAGCA 453

```

RESULT 5  
US-10-054-562a-6

; Sequence 6, Application US/10054562A

; Patent No. US20020165375A1

; GENERAL INFORMATION:

; APPLICANT: Chandrashekar, Ramaswamy

; APPLICANT: Morales, Tony H.

; TITLE OF INVENTION: PARASTIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES,

; FILE OF INVENTION: THEREOF

; CURRENT APPLICATION NUMBER: US/10/054,562A

; CURRENT FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: 09/812,642

; PRIOR FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: 09/323,427

; PRIOR FILING DATE: 1999-06-01

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 1372

; TYPE: DNA

; ORGANISM: Dirofilaria immitis

; US-10-054-562a-6

Query Match 43.4%; Score 504.2; DB 9; Length 1372;

Best Local Similarity 76.9%; Pred. No. 5.1e-122;

Matches 652; Conservative 0; Mismatches 193; Indels 3; Gaps 3;

```

OY 1 ATGATGATTCGCTTATTTGCTTTCTGCTACTACACTTATG-CATTGCTTATTCGATTCG 59
DB 32 ATATCATTTATTTCTCTGTTTATTTCTTACTTTCTACTTGACACTGATCATGATGCC 91
OY 60 GGTGACAATGGTGTGCAAGGTGAGCCAGAAATGGAATGTGAGCAACCTTCAATTAACAT 119
DB 92 TATGACAATGGTGTGCAAGGTGAGCCAGAAATGGAATGTGAGCGGAGCTTGTGATTAACAT 151
OY 120 CAATTTTATACAGTATGATTCGAAGGACATGTTATGTAAGAGTCTTTATGATCA 179
DB 152 CAATTTTATACAGTATGATTCGAAGGACATGTTATGTAAGAGTCTTTATGATCA 211

```

Query Match	43.4%	Score 504.2	DB 9	Length 1372
Best Local Similarity	76.9%	Pred. No.5.1e-122		
Matches 652	Conservative 0	Mismatch 193	Indels 3	Gaps 3
QY 1	ATGATGATGCTGCTTATTCCTTCTGTGACACCTATTG-CATTTGCTTATTCGATTC	59		
Db 1341	ATTATTCATTTATTCCTCTGTTATTCCTACTTTTCTACTGTAGCTACGTATTCATCGATCC	1282		
QY 60	GGTTGACAAATGCTGTCGAAGGTGAGCCAGAAATGGAATGTGACCACTTCATTAACAT	119		
Db 1281	TATTGACAAATGCTGTCGAAGGTGAGCCAGAAATGGAATGTGAGCGAGCTTCATTAACAT	1222		
QY 120	CAATTTAATACAGTAATGCAATTCAGAGACATGTTTATGTAAGGCTTTATGATCA	179		
Db 1221	CAATTTAATACAGTAATGCAATTCAGAGACATGTTTATGTAAGGAGCTCTATGATCA	1162		
QY 180	AGAAGTGTGCGGTAAATGATAAGGTGAGCGTCAGTTGCGGAAATTCCTTCATTTGA	239		
Db 1161	GGATGATATGTCGTTCCAGATAGTAAATGACCGCAGTAGCTGGAAATGGATTTGGCAATGA	1102		
QY 240	TTCAATGCAATGTTCGCGGTACAGATCTGGAATCCACGTGTAATTTGTTAACAACAC	299		
Db 1101	TTTCGTAAAGTTGGAAGATCAGATCCCTAAATCCCTCGGTGTTTGTGTAACAACCTG	1042		
QY 300	TGTTGTCATTTGCTTATTCATTCATTTATTTGTACCAAGTTGAATGCGATATTCAGTA	359		
Db 1041	AGTTGTCATTCATTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	982		
QY 360	ATGCTTTATACATGG-AAGCTGATAAACAGTTATGTC-ACAGATAGGATATCGAATC	417		
Db 981	ATGCTTTATATGGAAGATGATTAAGCGTTAGTACTGCTTCTTAATATCTGAAAG	922		
QY 418	ACAATGCTTTTCAAATCTCAAATGTCGAGTCCAGTATGCCGTTATGAAATTTTGAT	477		
Db 921	ACTACACATTCCAAATCTCAAAGTGTGACCAATGCCGTATGTCGATATGATGATTTGGA	862		
QY 478	GGTGACCAACCGGTCAACAGTTCAATTTGCTATCATTTGTCAGCGCATTTATCATTA	537		
Db 861	GGTGACCAACCGGTGTCACCTGTTTCATTTGCAATATATGAGATCATGATATATCACA	802		
QY 538	TGACATGCAATTCGTAACCGTGTGATCTTCTGCGCGGTGTCATTCCTGCTTGTGTC	597		
Db 801	TGACATGCTGATTCAGAGACTACAGATATCTGTCATTTACTATCATCATGATGTTG	742		
QY 598	GATATGCTGTAACGCTGATCTGTCGAAATTCCTAAATGCTGATGAGATGCTCTGTATTA	657		
Db 741	GATATGCAAAAGGTGATGTCAGTGAAGATTCCTAAATGGAAGAGATGCTCTTGACAA	682		
QY 658	TATTTGCTAATATATTTTGGAAATTCACAGATTTTATGCTGGCCAAAGCTCACGTA	717		
Db 681	TATTTACTACATTAATTTTGGAAATATATACAGATTTTATGCTGGCCAAAGCTCATGTT	622		
QY 718	TACAATATGCGGATCGATCACAGCTTTTCTATCAATGCGAGATAGTATACATTA	777		
Db 621	TATAAATATGAGATGATCATGAGAACTTTACTATCAATGCGAGATAGTATACATTA	562		
QY 778	GAAACCAATGCGAATGTGTCGACCAACAATGTTCCAGAACCAAGGATTTGGAAGTGT	837		
Db 561	GAGCCACATGCGAATGTGTCGACCAACAATGTCAGAGACCAAGGATTTGTCGCATA	502		
QY 838	AAAACAGG 845			
Db 501	AAATCTGG 494			

```

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: HW-8-2
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Brugia malayi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (158)..(892)
; OTHER INFORMATION:
US-10-054-562A-16
```

```

Query Match          39.7%; Score 461; DB 9; Length 892;
Best Local Similarity 78.1%; Pred. No. 8.2e-111;
Matches 554; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
```

```

QY 21 TTCTGTACTACACTATTGCTATGCTTATGATTCGCGTGACAAATGCTGCAAG 80
DB 184 TTGTGATATATGATATATGCAAGTATTAATGCTATTCATTAATGATGCTGCAAG 243
QY 81 TGAGCCAGAAATGAAATGAGTGGACCACTTCAATATCAATCAATTTTAATACAGTATGC 140
DB 244 TGAACTGAAATGAAATGAGTGGACCACTTCAATATCAATCAATTTTAATACAGTATGC 303
QY 141 ATTGAGAGCAATGTTTATGATGGAAGGCTTTATGATCAAGAAGGTTGCCGTAATGTA 200
DB 304 TTTTGAAGAGCAATGTTTATGATGGAAGGCTTTATGATCAAGAAGGTTGCCGTAATGTA 363
QY 201 AGGTGACGCTCAAGTGGCGGAATTTTCACTGATTTGATGATGATGATGATGATGATGAT 260
DB 364 AGGTGACGCTCAAGTGGCGGAATTTTCACTGATTTGATGATGATGATGATGATGATGATGAT 423
QY 261 AGCATCTGGAATCCACAGTGGTATTTTGTGAACAAACAGTGGTGGTGGTGGTGGTGGTGG 320
DB 424 AGCTTGTAAATCCACAGTGGTATTTTGTGAACAAACAGTGGTGGTGGTGGTGGTGGTGGTGG 483
QY 321 ATTATTTGTACCAAGTTGATGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 380
DB 484 ACAGTTTATCAAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
QY 381 TAAACAGTTAGTGCACAGATGAGTATCGAAATCAACAACGCTTTCAAACTCAAT 440
DB 544 TAAAGCTGTAGTGCACAAATTTGAAGTTCGCAAAATGACAAACCGTATTTGCTACACAAT 603
QY 441 TGTCCGATGCCAGTATGCGGTTATGAAATTTTGTGATGATGATGATGATGATGATGATGAT 500
DB 604 GGTACCAATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
QY 501 TCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
DB 664 CCAGTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
QY 561 TGATCTTCTGCGCGGTTGCTATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGAT 620
DB 724 TGATCTTCTGCGCGGTTGCTATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGAT 783
QY 621 GGAATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
DB 784 TAAATTTAATTAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
QY 681 TCCAAACAGATTTAATGCTGGCCAAAGAGCTCACGATATACAAATATGCG 729
DB 844 TCCAACTATTAATGCTGGCCAAAGAGCTCACGATATACAAATATGCG 892
```

```

RESULT 8
US-10-054-562a-18/c
; Sequence 18, Application US/10054562A
; Patent No. US20020165375A1
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES,
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Brugia malayi
US-10-054-562A-18
```

```

Query Match          39.7%; Score 461; DB 9; Length 892;
Best Local Similarity 78.1%; Pred. No. 8.2e-111;
Matches 554; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
```

```

QY 21 TTCTGTACTACACTATTGCTATGCTTATGATTCGCGTGACAAATGCTGCAAG 80
DB 709 TTGTGATATATGATATATGCAAGTATTAATGCTATTCATTAATGATGCTGCAAG 650
QY 81 TGAGCCAGAAATGAAATGAGTGGACCACTTCAATATCAATCAATTTTAATACAGTATGC 140
DB 649 TGAACTGAAATGAAATGAGTGGACCACTTCAATATCAATCAATTTTAATACAGTATGC 590
QY 141 ATTGAGAGCAATGTTTATGATGGAAGGCTTTATGATCAAGAAGGTTGCCGTAATGTA 200
DB 589 TTTTGAAGAGCAATGTTTATGATGGAAGGCTTTATGATCAAGAAGGTTGCCGTAATGTA 530
QY 201 AGGTGACGCTCAAGTGGCGGAATTTTCACTGATTTGATGATGATGATGATGATGATGATGAT 260
DB 529 AGGTGACGCTCAAGTGGCGGAATTTTCACTGATTTGATGATGATGATGATGATGATGATGAT 470
QY 261 AGCATCTGGAATCCACAGTGGTATTTTGTGAACAAACAGTGGTGGTGGTGGTGGTGGTGG 320
DB 469 AGCTTGTAAATCCACAGTGGTATTTTGTGAACAAACAGTGGTGGTGGTGGTGGTGGTGGTGG 410
QY 321 ATTATTTGTACCAAGTTGATGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 380
DB 409 ACAGTTTATCAAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 350
QY 381 TAAACAGTTAGTGCACAGATGAGTATCGAAATCAACAACGCTTTCAAACTCAAT 440
DB 349 TAAAGCTGTAGTGCACAAATTTGAAGTTCGCAAAATGACAAACCGTATTTGCTACACAAT 290
QY 441 TGTCCGATGCCAGTATGCGGTTATGAAATTTTGTGATGATGATGATGATGATGATGATGATGAT 500
DB 289 GGTACCAATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 230
QY 501 TCAATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
DB 229 CCAGTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 170
QY 561 TGATCTTCTGCGCGGTTGCTATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGAT 620
DB 169 TGATCTTCTGCGCGGTTGCTATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGAT 110
QY 621 GGAATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
DB 109 TAAATTTAATTAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 50
QY 681 TCCAAACAGATTTAATGCTGGCCAAAGAGCTCACGATATACAAATATGCG 729
```





Db 260 GATTGATCAAGACGATCAATTAATCTGATTAATACGGAGTGCAGTACCGGTTTAA 201  
QY 944 ACACCTTGAATAGGATGATTAATCAAGCTTGGCAGTTGAT 987  
Db 200 GCACGGTTGATATTAACGAGAGATCCGAACCTTCTCAGCAAAAT 157

## RESULT 11

US-09-938-842A-3231  
; Sequence 3231, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 3231  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3231

Query Match 3.5%; Score 40.2; DB 9; Length 2000;  
Best Local Similarity 45.2%; Pred. No. 2;  
Matches 147; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 148 GGACATGTTATGTGAAGAGCTTTATGATCAGAGAGTTCGGTATATGATGAAGGGA 207  
Db 819 GAAGATGTTAAGAAAGATCTATGCTCAGATGAAGATTCACAGTATTTGTAAGCA 878  
QY 208 CGCAGAGTCCGGAATTCATCTTCATTTGATTCATGATTCAGATTTGGCGTACAGCATCT 267  
Db 879 GCTTCAGGTTCTGAGATTTCTATTTGGGCTTTAGATTAAACACCTCTGTCACCTTG 938  
QY 268 CTGAATCCAGGTGATTTTGTGAACAACAAGTGTGTCATTTGCTTTCATTCATATTT 327  
Db 939 CTGATTCATTTCTCAGTTCTCAGTACAGACACTTAAAGTTTCGATTTTACTACTTTT 998  
QY 328 GTTACCAAGTTGATGTCATATTCGATTCAGATTCGATTTTACATGGAAGCTGATAAACA 387  
Db 999 ATTCTGATCTCAATCTATCTGACTTGAAGATTTTGTGTATACACAACAATATA 1058  
QY 388 GTTAGTGCACAGATTGAGTATGTAATCACAACCTGCTTTCAAACTCAATTTGCCG 447  
Db 1059 TGTAGTCAAGATCTACAAATTTCTCAATCATTTCAAGCTTCATCTCTCATGTATG 1118  
QY 448 ATGCCAGTATGCCGTTATGAATTT 472  
Db 1119 ATGGTAGAAGAACTGTGTACTT 1143

## RESULT 12

US-09-902-941-1883/c  
; Sequence 1883, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Matanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Panger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1883  
; LENGTH: 6799  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-1883

Query Match 3.3%; Score 38; DB 9; Length 6799;  
Best Local Similarity 59.1%; Pred. No. 15;  
Matches 65; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 235 TTTGATTCATGCAATGTTCGCGTACACGATCTGTAATCCAGTGTGTAATTTTGTAC 294  
Db 5002 TTTGTTACAGCAAAATCGAGTCCACGAGTCTCAATCGCATGTGTTCTCTGTCA 4943  
QY 295 ACAAGTGTGTCATTTGCTTTCATTCATTTATTTGTTACAAAGTTGATCG 344  
Db 4942 GTACTGTGTGCTATTTATTTCTCAGATGATTTATTTCTTTATTTATGG 4893

## RESULT 13

US-09-849-626-1883/c  
; Sequence 1883, Application US/09849626  
; Publication No. US20020197669A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya  
; APPLICANT: Panger, Gary  
; APPLICANT: Wang, Aijun  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Switzer, Anne  
; APPLICANT: McNeill, Patricia  
; APPLICANT: Clapper, Jonathan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C16  
; CURRENT APPLICATION NUMBER: US/09/849,626  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 1926  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1883  
; LENGTH: 6799  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-849-626-1883

Query Match 3.3%; Score 38; DB 9; Length 6799;  
Best Local Similarity 59.1%; Pred. No. 15;  
Matches 65; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 235 TTTGATTCATGCAATGTTCGCGTACACGATCTGTAATCCAGTGTGTAATTTTGTAC 294  
Db 5002 TTTGTTACAGCAAAATCGAGTCCACGAGTCTCAATCGCATGTGTTCTCTGTCA 4943  
QY 295 ACAAGTGTGTCATTTGCTTTCATTCATTTATTTGTTACAAAGTTGATCG 344  
Db 4942 GTACTGTGTGCTATTTATTTCTCAGATGATTTATTTCTTTATTTATGG 4893

## RESULT 14

US-10-017-754-1883/c

```
; Sequence 1883, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1883
; LENGTH: 6799
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-017-754-1883

Query Match
Best Local Similarity 59.1%; Score 38; DB 9; Length 6799;
Pred. No. 15;
Matches 65; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 235 TTTCATTCAGCATGTTGGCGGCGGACAGATCTGATCCATCCACGTCGATTTGTGTACA 294
DB 5002 TTTCATTCAGCATGTTGGCGGCGGACAGATCTGATCCATCCACGTCGATTTGTGTACA 294
DB 4942 GTACCTGTGCTGACTTATTTCCATGATGATTTATTCCTTTATTTTATG 4893

RESULT 15
; Sequence 18737, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Neomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18737
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035457.11
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
; OTHER INFORMATION: NT HIT: U01287.1, EVALU8 1.30e+00
; US-09-864-761-18737

Query Match
Best Local Similarity 46.3%; Score 36.2; DB 10; Length 510;
Pred. No. 10;
Matches 119; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 476 ATGTGACCAACCGGTCACCAAGTCATTTGCTATTCATTTGGTCAGCCAGTTATCAT 535
DB 414 ATGCTGATGACGAGCGCTGATCTGATGCAAGTATTTGGTGTGACAGATGTGATGAG 355
QY 536 AATGACATGCGATTTGGAACCGTTGATTTCTGCGCGGTTCATTCCTGCTTGG 595
DB 354 ATGTTGAAGCAGGTGTGAAGTGTGTGATGACAGGTGATGATGCTGTGATG 295
QY 596 TCGATGATGTAACGGTGTACTGTGCAAAATTTAAATGCTGATGATGCTCTTGATA 655
DB 294 ATGCTGATGTAAGTATGCTGTGCAAGTGTGATGCAAGTGTGAAAGAGTGTGATG 235
QY 656 AATATTTGCTAAATTAATTTGGAATATCCACAGATTTAATGCTGGCCCAAGAGCTCAG 715
DB 234 CAGGTGCTGAAGAAATGTTGATGCAAGTGAAGTGTGATGCTGATGATGCTGATG 175
QY 716 TATCAAAATATGCGGAT 732
DB 174 CTGGTAGTATGCTGCT 158

RESULT 16
; Sequence 5151, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
```

```
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/279,526
;; PRIOR FILING DATE: 2001-03-27
;; NUMBER OF SEQ ID NOS: 8481
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5151
;; LENGTH: 1103
;; TYPE: DNA
;; ORGANISM: Bacillus clausii
US-09-974-300-5151

Query Match          3.1%; Score 36.2; DB 10; Length 1103;
Best Local Similarity 56.2%; Pred. No. 16;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 471 TTGGATGGTGACCAACCGGTCAACCGATTGCTATCTATGTCAGCCAGTTTA 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 812 TTTCAGGATGTCCTTCAGTCTATTCGTTGTTGCTAGCCTTAGTGGCAATCGG 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 TCATAATGACATGCATCTCTGAACCGTTGATCTTTTCGCGGTTGTCATTCCTG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 872 TCATCGCCGATTTGCAATTTGCTCTGCTTTTGGCGCGGTGTCATCGCACCTG 931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 C 591
DB 932 C 932

RESULT 17
US-08-781-986A-33
; Sequence 33, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7563 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-33

Query Match          3.0%; Score 35.4; DB 7; Length 7563;
Best Local Similarity 53.2%; Pred. No. 79;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 16 ATTCCTTTCTACTACACTTAATGCAATGCTTAATTCGATTCGCCGTGCAATGCTGC 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5476 ATTGCTTACCGTATGAAGATGAAGTATTAATTTATTTATTAATCGCATGCAATGCGTGA 5535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 GAAGGTGAGCCCAATTAATGATGTGACCACTTCATAATCAATCAATTTAATACCGT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5536 AAAGTATGCTCGAAGAGATATGATTTGGTGGACAAAAGTAACGTTCATTCGCA 5595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 AATGCAATGCAAGACATGTT 156
DB 5596 CAAGCACGTTCTTCAATGCT 5616

RESULT 18
US-09-901-136-3
; Sequence 3, Application US/09901136
; Publication No. US20030039968A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1001273
; CURRENT APPLICATION NUMBER: US/09/901,136
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 378361
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(378361)
; OTHER INFORMATION: n = A,T,C or G
US-09-901-136-3

Query Match          3.0%; Score 35.4; DB 9; Length 378361;
Best Local Similarity 53.2%; Pred. No. 6-2e+02;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 119 TCATTTTAATACAGTAATGCAATTCGAAAGACATGTTATGTGAAGGCTTTATGATC 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279626 TCAGTATATAGTTAATATTAATTTTAAAGCATGTATAGAAAGATTTTCAGGTT 279685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 AAGAAGCTGCGGTATGATGAAGGTGAGCGTCAAGTCCGCAATTCACCTTCATTTG 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279686 ATGCAGGCTCTGATTTGCTGGAATAGAAAGAAATCACCTGTTTAAACATAAGATAG 279745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 ATTCAATGCAATGTCGCGTA 259
DB 279746 ATTTCACTGATGCTCAGTGAA 279766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 19
US-09-864-761-1160/C
; Sequence 1160, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
```

```

: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 1160
: LENGTH: 465
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC007389.2
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.89
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
: US-09-864-761-1160

Query Match      3.0%; Score 35.2; DB 10; Length 465;
Best Local Similarity 60.4%; Pred. No. 18;
Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Oy 59 CGGTGACATGCTGTGGAAGCTAGCCAGAAATGATGTGACCACTTCATACAA 118
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 CCGGGAAACAGATGAGAGTCTCTCAAAAATTAAGACTACGATGACCA 236

Oy 119 TCAATTTTATACAGTATGATTCATTCGAGACATG 154
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 GCAATTCACACTGCGGTATATATCAAGAAGATG 200

RESULT 20
: US-09-864-761-19488
: Sequence 19488, Application us/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
```

```

: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aomico-X-1
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 19488
: LENGTH: 660
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC008125.9
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.5
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.8
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3
: OTHER INFORMATION: NT HIT: AF044255.1, EVALUATE 1.10e-01
: US-09-864-761-19488

Query Match      3.0%; Score 35; DB 10; Length 660;
Best Local Similarity 46.2%; Pred. No. 25;
Matches 116; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

Oy 461 GTTGAATTTGGATGGAGGACCAACCGGTCAACAGTTTCATTTGCTATGTC 520
    || |||| | ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180

Oy 521 AGCCGTTTATCATTAATGACATCGATTCGTGAACCGTTGATACCTTCTGCGG 580
    | | | | | | | | | | | | | | | | | | | | | | | | | |
```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20772
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022334.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: AM992395.1, VALUE 9.50e-02
US-09-864-761-20772

Query Match          3.0%; Score 34.8; DB 10; Length 583;
Best Local Similarity 48.1%; Pred. No. 26;
Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 465 TGAATTTTGGATGTTGACCAACCGGTCAACGTTCAATTCCTTCATTCGTCAGCC 524
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 TGATGATGGGGATGATACACTGATGATGATGATGATGATGATGATGATGATGATG 315
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 AGTTATCATTAATGACATGCGATTCGAAACCGTTCGATCTTCGCGCGGTGACCA 584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 CGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 TTCCTGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 644
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 376 TGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 TGCCTTTCATTAATTAATTTGCTAAATA 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 436 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 461
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 23
US-09-815-242-3929
; Sequence 3929, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3929
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-815-242-3929

Query Match          3.0%; Score 34.8; DB 10; Length 1608;
Best Local Similarity 53.7%; Pred. No. 47;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 601 GATGTAACGGTGTACTGTGGAATTCCTAATCTGATGATGATGATGATGATGATGAT 660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 910 GATCCATATATTCTGTGTGAAGCATTAACAATGCTGCTTGAATTTGATTCAGAT 969
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 TTGCTAATTAATTTGATATCAACAGATTATGCTGCGCAAGAAGCTCAGCTATAC 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 970 ATCGAATTCATTTGGTGTATTCACAAAGATTACTGCTGAATGTCGCGTGAAGAATT 1029
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 AATATCGCATCG 734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1030 GCTTCACGATG 1043
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 24
US-09-815-242-6630
; Sequence 6630, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```



```

      PRIOR APPLICATION NUMBER: US 09/608,408
      PRIOR FILING DATE: 2000-06-30
      PRIOR APPLICATION NUMBER: US 09/774,203
      PRIOR FILING DATE: 2001-01-29
      NUMBER OF SEQ ID NOS: 49117
      SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
      SEQ ID NO 4012
      LENGTH: 1959
      TYPE: DNA
      ORGANISM: Homo sapiens
      FEATURE:
      OTHER INFORMATION: MAP TO AL022334.1
      OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL = 1.9
      OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
      OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
      OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.3
      OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
      OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
      OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
      OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
      OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
      OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
      US-09-864-761-4012
    Query Match          3.0%; Score 34.8; DB 10; Length 1959;
    Best Local Similarity 48.1%; Pred. No. 52;
    Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

      465 TGAATTTGGATGGTGGACCAACCGGTCACACGTTCAATTGCTATCATGCTCAGCC 524
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Db 761 TGATGATGGGATGATGTCATGCACAGTGAAGTGGTATGCTGATGATGATGATGA 820

      525 AGTTATCATTAATGAGACATCGATTCGTAACCGTTGATCTTCCGCGGTTGTCCA 584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Db 821 CGGGATATATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 880

      585 TTTCCTGTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 644
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Db 881 TGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      Db 941 TGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

    RESULT 26
    US-10-210-296-1
      Sequence 1, Application US/10210296
      Publication No. US20030021802A1
      GENERAL INFORMATION:
      APPLICANT: Pfizer Products Inc.
      TITLE OF INVENTION: LAWSONIA INTRACELLULARIS PROTEINS, AND RELATED
      TITLE OF INVENTION: METHODS
      TITLE OF INVENTION: AND MATERIALS
      FILE REFERENCE: PC10589A
      CURRENT APPLICATION NUMBER: US/10/210,296
      CURRENT FILING DATE: 2002-08-01
      PRIOR APPLICATION NUMBER: US/09/669,065
      PRIOR FILING DATE: 2000-10-12
      NUMBER OF SEQ ID NOS: 102
      SOFTWARE: PatentIn Ver. 2.1
      SEQ ID NO 1
      LENGTH: 6617
      TYPE: DNA
      ORGANISM: Lawsonia intracellularis
      US-10-210-296-1
    Query Match          3.0%; Score 34.8; DB 9; Length 6617;
    Best Local Similarity 50.6%; Pred. No. 1e+02;
    Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

      87 AGAATTGAATGTGGACCACTTCATTAACATCAATTTTAATACAGCTAATGATTCGCA 146
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 1490 AGAATTGCTAATGATTTTCAGAACTTAATGATTCATGATCAAGCTGCGTTTGA 1549  
QY 147 AGCAGATGTTATGTAAGGCTTTATGATCAGAAGCTGCCGATATGTAAGGCG 206  
Db 1550 AGAACAAGTTCTTGAGAAAGCAGCTGAGACGAAGACATGCCATGATGAAGATTA 1609  
QY 207 ACCTCAAGTTGCCGAATTTTCATTTCCATTTGATTCATGATGATGTT 252  
Db 1610 TCTTCGTCGATTAAGATATGAAATGCCACGACGACGAGGAGAGCT 1655

## RESULT 27

US-09-070-927A-217/c  
; Sequence 217, Application US/09070927A  
; Patent No. US20020120116A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Kunsch  
Patrick J. Dillon  
Steven Barash  
TITLE OF INVENTION: Enterococcus faecialis polynucleotides and polypeptides  
NUMBER OF SEQUENCES: 982  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070.927A  
FILING DATE: 04-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8395 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-09-070-927A-217

Query Match 3.0%; Score 34.8; DB 10; Length 8395;  
Best Local Similarity 53.7%; Pred. No. 1.2e+02;  
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 601 GATCGTAACGGTGAATCTGGAATTTCTAAATGCTGATGATGCTCTTGATTAATAT 660  
Db 4247 GATCCATATATTTCTGTTGAGAACATTAACATGCTGCTGTTGATTTGATTCACAT 4188  
QY 661 TTGCTAAATATTTGGAATATCCACAGATTTAATGCTGCGCAAGAAGCTCAGTATAC 720  
Db 4187 ATCAAAATGTATGGTTGATTCACAAAGATTAACCTGCTGAATAATGTCGAAAGATT 4128  
QY 721 AATATGCGGATCG 734

Db 4127 GGTTCAGCAGATCG 4114

## RESULT 28

US-09-070-927A-405  
; Sequence 405, Application US/09070927A  
; Patent No. US20020120116A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Kunsch  
Patrick J. Dillon  
Steven Barash  
TITLE OF INVENTION: Enterococcus faecialis polynucleotides and polypeptides  
NUMBER OF SEQUENCES: 982  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070.927A  
FILING DATE: 04-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 405:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1479 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 405:  
US-09-070-927A-405

Query Match 3.0%; Score 34.6; DB 10; Length 1479;  
Best Local Similarity 44.3%; Pred. No. 50;  
Matches 136; Conservative 1; Mismatches 170; Indels 0; Gaps 0;

QY 111 AATAACAATCAATTTTAATACGTAATGATTCGAGAGACATGTTATGTAAGGCT 170  
Db 1061 AAGAACAATTTATTTATGAAATTTAGCAATATGAATAATGACATGCAAGAAATGT 1120  
QY 171 TTATCATCAAGAGGTTCCGTAATGATGACAGTGCAGTCAAGTGGCCGAATTTCACT 230  
Db 1121 TGAAGCTAAGATGCTGTTATTTATGTAAGCTTAGTAAATCGAAGCAGTCTTATC 1180  
QY 231 TCCATTTGATTCATGCAATGTTCCGCTACAGATCTGATCCACGCTGATTTTGT 290  
Db 1181 AAACAAGATCAATATGCCAATGAATTTTACCAACCGCATGTCGRTTAAGTTAGCT 1240  
QY 291 AACACAACGTGTGATTTGCTTTCATCATCTTATTTGTAACAAAGTTGATCGTCATA 350  
Db 1241 ATCAGCTGTGAAAACATTCAAAAGGCCACAGTGTTTTGAAGCCGTAGTCATCCAGA 1300



RESULT 33  
US-10-054-562A-12/c  
; Sequence 12, Application US/10054562A  
; Patent No. US20020165375A1

```

RESULT 34
US-09-864-761-25339
; Sequence 25339, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25339
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023000.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: NT HIT: AF164151.1, EVALUE 6.80e-01
; OTHER INFORMATION: EST_HUMAN HIT: BE336733.1, EVALUE 3.20e+00
US-09-864-761-25339
```

```
Query Match          2.9%  Score 34;  DB 10;  Length 325;
Best Local Similarity 56.1%;  Pred. No. 30;
Matches 64;  Conservative 0;  Mismatches 50;  Indels 0;  Gaps 0;
```

```
QY 547 GATTCTGAACCGTTGACTTCTTCGCGGGTTCATTCGCTTGTGCGATGATG 606
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 124 GATTATCATATATGTATGATGATGATGATGATGATGATGATGATGATGATG 183
QY 607 AACGGTGAACCTGTGGAATCTTAAATGCTGATGATGATGATGATGATGAT 660
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 184 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
```

```

RESULT 35
US-09-864-761-8619
; Sequence 8619, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8619
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023000.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-8619
```

```
Query Match          2.9%  Score 34;  DB 10;  Length 556;
Best Local Similarity 56.1%;  Pred. No. 41;
Matches 64;  Conservative 0;  Mismatches 50;  Indels 0;  Gaps 0;
```

```
QY 547 GATTCTGAACCGTTGACTTCTTCGCGGGTTCATTCGCTTGTGCGATGATG 606
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 158 GATTATCATATATGTATGATGATGATGATGATGATGATGATGATGATGATG 217
QY 607 AACGGTGAACCTGTGGAATCTTAAATGCTGATGATGATGATGATGATGAT 660
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 218 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 271
```

```

RESULT 36
US-09-745-288-91
; Sequence 91, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jlangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 895
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-91
```

```
Query Match          2.9%  Score 34;  DB 10;  Length 895;
Best Local Similarity 54.9%;  Pred. No. 54;
Matches 67;  Conservative 0;  Mismatches 55;  Indels 0;  Gaps 0;
```

```
QY 282 TATTTTGTAAACACACAGTGTGCTATTCGTTTCATTCATATATTTGTAACAAAGTTGA 341
      |||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
DB 150 TATATCCTTAATCTCAATATATGTCACCTTATTAATAAAGGATTTGGTAAACACATGCAT 209
QY 342 TCGTCATATGATGATCAATGCTTTTACATGGAAGCTGATAAACAATTAGTGACACAGAT 401
```

Db 210 TTGTGACACAGATGCCAAAATTTATACAGTAGTAATGACCAACCAAGAGATATACAC 269  
402 TG 403  
270 TG 271

## RESULT 37

US-09-764-872-725/C  
; Sequence 725, Application US/09764872  
; Publication No. US20030050231A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA125  
; CURRENT APPLICATION NUMBER: US/09/764,872  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 957  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 725  
; LENGTH: 2322  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-872-725

Query Match 2.9%; Score 34; DB 9; Length 2322;  
Best Local Similarity 54.9%; Pred. No. 94;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 282 TATTTTGTACACACAGCTGTGTCATTGCTTTCATCCATTAATTTTACCAAGTTGA 341  
Db 2179 TATATCTTACTTCACTTATTTGTCACCTTATTAATTAAGGTGATTCCTTAACACATGCAT 2120  
QY 342 TCGTGCATATCGAGTACATGCTTTTACATGAGAGCTGATTAACAGTTAGTCACAGAT 401  
Db 2119 TTGTGACACAGATGCCAAAATTTATACAGTAGTAATGACCAACCAAGAGATATACAC 2060

QY 402 TG 403  
Db 2059 TG 2058

## RESULT 38

US-08-781-986A-272/C  
; Sequence 272, Application US/08781986A  
; Publication No. US20030054436A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob

; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 272:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-781-986A-272

Query Match 2.9%; Score 34; DB 7; Length 6035;  
Best Local Similarity 50.6%; Pred. No. 1.6e+02;  
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 237 TGATTATGCAATGTCGCGCTACACGATCTCGAATCCAGTGTATTTTGTACAAAC 296  
Db 5958 TTGTATGCAATGATCGCGCTAAGCTCAACAATGATATTTATAGTTGTACAGC 5899  
QY 297 AACTGTGTCAATTCGTTTCATCCATTAATTTGTACAAAGTTGATGCTGATTCGACT 356  
Db 5898 AGCATTTAGATATTAATTAATTTATACCTGTCACAGTCAACAAGAAATTTATTTGCTGTCCC 5839  
QY 357 ACATGCTTTTACATGCAAGCTGATAAAGAGTTAGTCACA 398  
Db 5838 TGCGCGCTTTGACATATACGACCTCAAAATTAATTCACA 5797

## RESULT 39

US-09-815-242-4340  
; Sequence 4340, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4340  
; LENGTH: 1170  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-4340

Query Match 2.9%; Score 33.8; DB 10; Length 1170;  
Best Local Similarity 52.5%; Pred. No. 71;  
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Search completed: March 28, 2003, 17:58:31  
Job time : 529 secs

```
OY 16 ATGCTTCTGTACTACACTTATTCATGCTTTATTCGATCCGGTTGACAAATGGTGTC 75
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 796 ATGCTTACCGTATGAAGATATGAACTTTTATTATTACCGATCAATGCGTGCA 855
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

OY 76 GAAGTGAGCCGAATTAATGATGGACCACTTCATACAAATTTTAATACAGT 135
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 856 AAAGGTATGCCGTGAAGAGATATGATTTGGGTGTCAAAAGTAAGTCAATCGCAA 915
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

OY 136 AATGCATTCGAAGACATGTT 156
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 916 CAAGCACGCTTTCGAATATGT 936
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

## RESULT 40

```
US-09-815-242-8204
: Sequence 8204, Application US/09815242.
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlssen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Cair, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011a
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8204
: LENGTH: 1182
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1182)
US-09-815-242-8204
```

## Query Match

2.9%; Score 33.8; DB 10; Length 1182;  
Best Local Similarity 52.5%; Pred. No. 72;  
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```
OY 16 ATGCTTCTGTACTACACTTATTCATGCTTTATTCGATCCGGTTGACAAATGGTGTC 75
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 805 ATGCTTACCGTATGAAGATATGAACTTTTATTATTACCGATCAATGCGTGCA 864
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

OY 76 GAAGTGAGCCGAATTAATGATGGACCACTTCATACAAATTTTAATACAGT 135
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 865 AAAGGTATGCCGTGAAGAGATATGATTTGGGTGTCAAAAGTAAGTCAATCGCAA 924
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

OY 136 AATGCATTCGAAGACATGTT 156
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 925 CAAGCACGCTTTCGAATATGT 945
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 15:14:55 ; Search time 1642 Seconds  
(Without alignments)  
11451.255 Million cell updates/sec

Title: US-10-054-562a-3

Perfect score: 1161

Sequence: 1 atgatgatctctatctatgc.....ttcgtccaatcagaagca 1161

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inu:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	481.8	41.5	721	12	BG310598	BG310598 SMOV3MCA1
2	445.2	38.3	638	12	BF482157	BF482157 SMOV3MCA1
3	432.6	37.3	775	12	BF824707	BF824707 SMOV3MCA1
4	403	34.7	758	12	BF918213	BF918213 SMOV3MCA1
5	392.6	33.8	647	12	BF727586	BF727586 SMOV3MCA1
6	390.2	33.6	665	9	A1111196	A1111196 SMOV3MCA1

7	387.2	33.4	740	12	BF727527	BF727527 SMOV3MCA1
8	385	33.2	648	12	BF400424	BF400424 SMOV3MCA1
9	376.8	32.5	646	13	BM285148	BM285148 KH95C07.Y
10	376.2	32.4	747	12	BG809012	BG809012 SMOV3MCA1
11	372.2	32.1	628	9	AA701731	AA701731 SMOV3MCA1
12	370.2	31.9	644	12	BF400417	BF400417 SMOV3MCA1
13	369.8	31.9	774	12	BF599177	BF599177 SMOV3MCA1
14	364.2	31.4	672	9	AA618895	AA618895 SMOV3MCA1
15	363.6	31.3	612	13	BM284858	BM284858 KH95H05.Y
16	358.6	30.9	639	9	A1322117	A1322117 SMOV3MCA1
17	355.8	30.6	610	9	AA688077	AA688077 SMOV3MCA1
18	351.6	30.3	648	13	BM284957	BM284957 KH93A07.Y
19	336.8	29.0	635	12	BF482126	BF482126 SMOV3MCA1
20	333.8	28.8	624	13	BM284979	BM284979 KH93C09.Y
21	331.4	28.5	554	12	BF482099	BF482099 SMOV3MCA1
22	324.2	27.9	776	9	AA585626	AA585626 SM29CA34
23	312.4	26.9	537	9	AA688051	AA688051 SMOV3MCA1
24	307	26.4	691	12	BF727629	BF727629 SMOV3MCA1
25	306.8	26.4	552	9	A1322068	A1322068 SMOV3MCA1
26	299.6	25.8	615	9	A1317885	A1317885 SMOV3MCA1
27	299	25.8	543	12	BF758436	BF758436 SMOV3MCA1
28	297.6	25.6	694	10	AM355791	AM355791 SMOV3MCA1
29	292.6	25.2	534	13	BI501509	BI501509 TM01C11.Y
30	291.2	25.1	592	13	BI783374	BI783374 KH19B11.Y
31	285	24.5	592	13	BF400320	BF400320 SMOV3MCA1
32	280.4	24.2	622	12	BF400320	BF400320 SMOV3MCA1
33	278.8	24.0	363	9	AA625020	AA625020 SMOV3MCA1
34	277.8	23.9	558	10	AM313040	AM313040 SMOV3MCA1
35	275.6	23.7	543	10	AM312995	AM312995 SMOV3MCA1
36	274.6	23.7	577	10	AA841200	AA841200 MB3D6AA4G
37	269.2	23.2	705	12	BF918256	BF918256 SMOV3MCA1
38	269.2	23.2	466	9	AA625024	AA625024 SMOV3MCA1
39	267.8	23.1	459	9	AA618952	AA618952 SMOV3MCA1
40	263	22.7	459	9	AA625010	AA625010 SMOV3MCA1
41	260.4	22.4	395	9	AA625010	AA625010 SMOV3MCA1
42	257.6	22.2	523	13	BI781683	BI781683 KH15H01.Y
43	257	22.1	369	9	AA625022	AA625022 SMOV3MCA1
44	253.6	21.8	492	10	AM313041	AM313041 SMOV3MCA1
45	249	21.4	541	13	BI662949	BI662949 TM14E03.Y

#### ALIGNMENTS

RESULT 1  
BG310598  
LOCUS 721 bp mRNA linear EST 23-FEB-2001  
DEFINITION SMOV3MCA15F02SK Onchocerca volvulus molting L3 larva cDNA  
(SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA15F02 5',  
mRNA sequence.

ACCESSION BG310598  
VERSION BG310598.1 GI:13112315  
KEYWORDS EST.  
SOURCE Onchocerca volvulus.  
ORGANISM Onchocerca volvulus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Onchocerca.

REFERENCE 1 (bases 1 to 721)  
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Taney,S. and Lustigman,S.  
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
JOURNAL Unpublished (1997)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu

FEATURES  
source  
1..721  
/organism="Onchocerca volvulus"

```
/strain="Kumba, Camerouns"  
/db_xref="taxon:6282"  
/clone="SMOV3MCAM48D12SK"  
/clone.lib="Onchocerca volvulus molting L3 larva cdna  
(SL96MLM-Ovml3)"  
/dev_stage="molting L3"  
/lab_host="XLI-Blue MRF"  
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:  
Xho I; Filarial nematode parasite of humans. Third-stage  
larvae, L3, were isolated from infected black flies in  
Cameron (forest strain). The L3 were cultured in 20% FCS  
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in  
culture. L3 of O. volvulus molt to fourth-stage larvae by  
day 5 in culture. mRNA was isolated from approximately  
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3  
in culture, and converted to double-stranded cDNA using  
reverse transcriptase and oligo(dT) followed by RNase H  
and DNA pol I. The library was constructed in the lambda  
Uni-Zap XR vector and has 1 x 10E6 independent  
recombinants and the average insert size is ~1200 bp. The  
library was constructed by Sara Lustigman and Michelle  
Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.  
The library is available from Dr. Sara Lustigman (email:  
slustigmen@nybc.org)."  
BASE COUNT      213 a      140 c      158 g      210 t  
ORIGIN
```

```
Query Match      41.5%; Score 481.8; DB 12; Length 721;  
Best Local Similarity 85.4%; Pred. No. 1.5e-123;  
Matches 537; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```
OY 160 GTGAAGGCTTTATGATCAGAAGGTTGCCGTAATGATGAAGGAGCGTCAAGTGGCC 219  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 1 GTGAAGGCTTTATGATCAGAAGGTTGCCGTAATGATGAAGGAGCGTCAAGTGGCC 60  
OY 220 GGAATTCCTCCATTCATTCATGCAATGTTGCCGCTACAGATCTGTAATCCAGT 279  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 61 GGATTAAGCTTCGTTGATTCCTGTAATGAGCAGTACAGTTCGTTAAATCCAGCT 120  
OY 280 GGTATTTTGTAAACAACAACCTGTTGTCATTCGTTTCATTCATTCATTCATTCATTC 339  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 121 GGTATTTTGTACATCATCATGTTGTCATCTCATTCATTCATTCATTCATTCATTCAT 180  
OY 340 GATCGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 399  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 181 GATCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
OY 400 ATTGAGTATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 459  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 241 CTTGAAGTTTCCGAATGAGCAACTGATTTGCAACGAATTTGATGATGATGATGATGATG 300  
OY 460 CTTATGAATTTTGGATGGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 519  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 301 CGATATGATATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
OY 520 CAGCCAGTTATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 579  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 361 CAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
OY 580 GTTCATTCCTGCTGTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 421 GTACATTCATGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
OY 640 GGATGCTGCTGATTAATATTTGCAATTTTGGATTAATTCACACAGATTAATGGGT 699  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 481 GGTGCTGTTGGACAAATATTTGCTCAACAACCTTGGATTAATTCACAGATTAATGGGT 540  
OY 700 GGCCAGAAGGCTCAGATTAATTAATGCGATGATGATGATGATGATGATGATGATGATGAT 759  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 541 GGACAGAAGGCTCAGATTAATTAATGCGATGATGATGATGATGATGATGATGATGATGAT 600  
OY 760 ATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 788  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB 601 ATTAGTATTACATTAAGAACCACCAAG 629  
RESULT 2  
BF482157  
LOCUS  
DEFINITION  
BF482157  
SMOV3MCAM48D12SK Onchocerca volvulus molting L3 larva cdna  
(SL96MLM-Ovml3) Onchocerca volvulus cdna clone SMOV3MCAM48D12 5',  
mRNA sequence.  
ACCESSION  
BF482157  
VERSION  
BF482157.1 GI:11565458  
KEYWORDS  
EST.  
SOURCE  
Onchocerca volvulus.  
Onchocerca volvulus.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Splirrida; Filarioidea;  
Onchocercidae; Onchocerca.  
REFERENCE  
1 (bases 1 to 638)  
Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
Genes expressed in molting L3 larvae of Onchocerca volvulus  
Unpublished (1997)  
JOURNAL  
COMMENT  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: pluescript SK.  
location/Qualifiers  
1..638  
/organism="Onchocerca volvulus"  
/strain="Kumba, Camerouns"  
/db_xref="taxon:6282"  
/clone="SMOV3MCAM48D12"  
/clone.lib="Onchocerca volvulus molting L3 larva cdna  
(SL96MLM-Ovml3)"  
/dev_stage="molting L3"  
/lab_host="XLI-Blue MRF"  
/note="Vector: lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:  
Xho I; Filarial nematode parasite of humans. Third-stage  
larvae, L3, were isolated from infected black flies in  
Cameron (forest strain). The L3 were cultured in 20% FCS  
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in  
culture. L3 of O. volvulus molt to fourth-stage larvae by  
day 5 in culture. mRNA was isolated from approximately  
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3  
in culture, and converted to double-stranded cDNA using  
reverse transcriptase and oligo(dT) followed by RNase H  
and DNA pol I. The library was constructed in the lambda  
Uni-Zap XR vector and has 1 x 10E6 independent  
recombinants and the average insert size is ~1200 bp. The  
library was constructed by Sara Lustigman and Michelle  
Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.  
The library is available from Dr. Sara Lustigman (email:  
slustigmen@nybc.org)."
```

```
BASE COUNT      188 a      123 c      140 g      187 t  
ORIGIN  
Query Match      38.3%; Score 445.2; DB 12; Length 638;  
Best Local Similarity 83.1%; Pred. No. 2.5e-113;  
Matches 507; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
OY 257 GACACGATCTCTGATATTCAGTGTATTTTGTAAACAACAACGTTCTCATTTGCTTC 316  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 3 GCACGAGTTGCTGTTAAATCCCGGTGATTTTGTACATTCATTCATTCATTCATTCATTC 62  
OY 317 ATCCATTTTGTGTTACCAAGTTGATCGATATCGATGATGATGATGATGATGATGATGAT 376  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 63 ATCCATTTTGTGTTGCAAAAGTTGATCGATATCGATGATGATGATGATGATGATGATGAT 122  
OY 377 CTGATTAACAAGTTAGTGCACAGATTTGAGGTATCGAATTAACAACGCTTTTCAAACTG 436  
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 123 CTGATAAGACAGTTAGGCGCTCAACTTGAAGTTCCGAATGACAACTGCATTGGCAAC 182
QY 437 AATTGCGCGATGCCAGTATGCGCTTATGAATTTGGATGGTGACCAACGGTCAAC 496
Db 183 AATTGCGCGATGCCAGTATGCGCTTATGAATTTGGATGGTGACCAACGGTCAAC 242
QY 497 CAGTTCAATTTGGTATCATGTCAGCAGTATTCATTAATGACATGCGATTCTGAAA 556
Db 243 CTGTTCAATTTGGTATCATGTCAGCAGTATTCATTAATGACATGCGATTCTGAAA 302
QY 557 CCGTTGATCTTCTGCGCGGTTGTCATTCGCTTGTGCGATGATGATGCGTGATA 616
Db 303 CAGTTGATCTTCTGCGCGGTTGTCATTCGCTTGTGCGATGATGATGCGTGATA 362
QY 617 CTGTGAAATTTCAATAGCTGATGATGCTCTGTGATTAATATTGCTTAATATTGG 676
Db 363 AGGTGAAATTTCAATAGCTGATGATGCTCTGTGATTAATATTGCTTAATATTGG 422
QY 677 AATATCCACAGATTTATGCTGCGCAAGAGTCAAGTATACAAATATGCGGATCGAT 736
Db 423 AATATCCACAGTATTTATGCTGCGCAAGAGTCAAGTATACAAATATGCGGATCGAT 482
QY 737 CACAGCTTTCTATCAATGCGCAAGTATACAAATTAAGAACCAATAGCGAATGTG 796
Db 483 CACAGCTTTCTATCAATGCGCAAGTATACAAATTAAGAACCAATAGCGAATGTG 542
QY 797 TTGGACCAAGATTTGCAAGCAAGATTCGAGAGTGTAAACAGGTGGTCCGAC 856
Db 543 CAGGACCAAGATTTGCAAGCAAGATTTGAGAGTGTGTCGCAAGATTTGGTGTG 602
QY 857 CAAACCTGC 866
Db 603 CTCGAGCTGC 612

RESULT 3
BF824707 775 bp mRNA linear Est 13-JAN-2001
LOCUS SMOV3MCA52G11SK Onchocerca volvulus molting L3 larva cDNA
DEFINITION (SI96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA52G11 5',
mRNA sequence.
ACCESSION BF824707 GI:1216567
VERSION BF824707.1
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 775)
AUTHORS Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL Unpublished (1997)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genomesmith.edu
Seq primer: pluescript SK.
Location/Qualifiers
1..775
/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone_lib="SMOV3MCA52G11"
/clone_lib="Onchocerca volvulus molting L3 larva cDNA
(SI96MLM-Ovml3)"
/dev_stage="molting L3"
/lab_host="XLI-Blue MRF"
/Note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
```

```
larvae, L3, were isolated from infected black flies in
Cameron (forest strain). The L3 were cultured in 20% FCS
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNA pol I. The library was constructed in the Lambda
Uni-Zap vector and has 1 x 106 independent
recombinants and the average insert size is ~1200 bp. The
library was constructed by Sara Lustigman and Michelle
Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
The library is available from Dr. Sara Lustigman (email:
slustigmen@mc.org)."
BASE COUNT 215 a 144 c 166 g 250 t
ORIGIN
Query Match 37.3%; Score 432.6; DB 12; Length 775;
Best Local Similarity 82.4%; Pred. No. 8.8e-110;
Matches 556; Conservative 0; Mismatches 114; Indels 5; Gaps 5;
QY 50 ATTGATTCGCGTTGACAAATGCTGCAAGGTGAGCCAGAAATGAATGGACCAACTT 109
Db 106 ATGCTATTCGCGTTGATTAAGCGGTAGAGGTGAACCAAAATGAATGGTCCCAACT 165
QY 110 CAATTAACATCAATTTTATATACAGTAAATGATTCGAAGCAAGATTATGTGAAGTTC 169
Db 166 CAATTAAGTGCACAACTTAATACCTGTAATCATTGGAAGGACAAATGATACCTGAAAGCT 225
QY 170 TTTATGATCAAGAAGGTTGCGTAAATGATGAGGTGACGTCAGTATTCGCGGATTTTCAAC 229
Db 226 TATACGATCAGCGAGATGCCGAATGATGAGGTGACGTCAGTATTCGCGGATTTTCAAC 285
QY 230 TTCATTGATTCATGCAATGTCGCGTACACAGATCTGTAATCCACGTGATATTTTG 289
Db 286 TTCGTTGATTCCTGTAATGATACAGTACAGTACAGTCTGTAATCCACGTGATATTTTG 345
QY 290 TAAACAACGTCGTGATTCGTTGATTCATTAATTTGTTACCAAGTGAATGTCGTCAT 349
Db 346 TCACATCAGTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 405
QY 350 ATCAGTACATGCTTTTACATGGAAGCTGATTAACAGTATGACAGATGAGGTAT 409
Db 406 ATGCAATACATGCTTTTACATGGAAGCTGATTAACAGTATGACAGATGAGGTAT 465
QY 410 CTGAATACCAACTGCTTTTCAAACTCAATTTGCCGATGCCAGTATGCCGTTATGAAA 469
Db 466 CCGAAATGACAACTGCAATTTGCAACACAAATTTG -CCGATGCCGATGCCGATGAGA 524
QY 470 TTTTGGATGGTGACCAACCGGTAAACAGTTCATTTGCTATATGTCATGACGACTTT 529
Db 525 TTCTTGATGGTGACCAACCGGTAAACAGTTCATTTGCTATATGTCATGACGACTTT 584
QY 530 ATCATTAATGACATGCGATTTCTGAACCGGTGATTTCTTCTGCGCGGTTCCATTCCT 589
Db 585 -CCATTAATGACATGCGATTTCTGAACCGGTGATTTCTTCTGCGCGGTTCCATTCAT 642
QY 590 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
Db 643 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701
QY 650 TTGATTAATTTGCTTAATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 709
Db 702 TGGGACAAATTTTCTCAACCTGGGAATATTC -ACTGATTTAATGGCTGACAGAAAG 760
QY 710 CTCACGTATCAAAAT 724
Db 761 CTCACGTATCAAAAT 775

RESULT 4
BF918213
```



The library is available from Dr. Sara Lustigian (email: [slustigian@bc.rockefeller.edu](mailto:slustigian@bc.rockefeller.edu))

BASE COUNT	186 a	144 c	134 g	183
ORIGIN				

Query Match	33.8%;	Score 392.6;	DB 12;	Length 647;
Best Local Similarity	84.8%;	Pred. NO. 1.2e-98;		
Matches 452;	Conservative 0;	Mismatches 79;	Indels 2;	Gaps 1;

9y 1 ATGAGTATTCGCTTATTCGTTTCGTACTACCTTATTCGATTCG 66  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 84 ATGAGCTTCGCTTATTCGCTTTCGACATCGGCTTTTGACATGCTATTCATTCCT 14

Qy 61 GTTGACATGGTGTGAGAGTGAGCCAGAATTGAATGTGACCACCTTCATAAACATC 120  
|||||  
Db 144 GTTGACAACGGTGTCTGGAAGAGAGGCCAGAAATGAGTGGCGACCACTTCGATAACTATC 200

09 121 AATTAAATACACGTAATGCATTGCAAGGACATGTTATGTGAAGGCTCTTATGATCAA 18  
 Db 204 AATTTTAATACACGCAATGCATTGCAAGGACATGTTATGTGAAGGCTCTTACGATCAA 26

07 GAAAGTGGCCGATGAAGAAGGTGGACCTCAAGTTGCCGGAATTCACTTCATTGGAT 24  
161 | | | | | | | | | | | | | | | | | |  
Db 264 GCAGGTGCCCGAATGATGAAGTGACGTCAGTTGCCGGAATTCACTTCCCCTCGAT 32

Db 324 TCTTGCATGTTGCACGCACAGATCTCTGATCCGCGAGGTACTTTGTACACACTACT 387

Db 384 GTTGATATCTCCTTCACCCACTGTTGTGACAAAGTTGATCGTGCATACCGAGTACAG 444

Db 444 TGCTTTACATGGAAGCTGACAAAACGGTGCACCCAGATCGAAGTGTCCGAAATCACA 50

Db 504 ACTGCTTTCAAACTCAGATTGTTCCATGCGTGTTCGCCGATATGAATCTTGATGGT 564

Db 564 GGACACTGGACAAACCAATTCAGTTGGCTACCATTCGGGGCAGCCAGTTTAT 616

RESULT 6	LOCUS	665 bp	MRNA	linear	EST 31-AD
AI111196	AI111196	665 bp	MRNA	linear	EST 31-AD

(SI96MLM-OVML3) *Onchocerca volvulus* cDNA SMOV3MCA1232  
 mRNA sequence.  
 A111196  
 A111196.1 C1.3510080  
 ACCESSION  
 VERSION

KEYWORDS	EST
SOURCE	<i>Onchocerca volvulus</i> .
ORGANISM	<i>Onchocerca volvulus</i>
Phylum:	Metazoa
Class:	Nematoda
Order:	Chromadoro
Family:	Onchocercidae
Genus:	<i>Onchocerca</i>
Species:	<i>O. volvulus</i>

REFERENCE  
AUTHORS  
TITLE

Uncinocercariae; Uncinocerca.  
1 (bases 1 to 665)  
Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
Genes expressed in molting 13 larvae of *Uncinocerca volvulus*

JOURNAL Unpublished (1997)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology

Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: plusescript SK.  
Location: Quailfliers

FEATURES	Location/Qualifiers
source	1. .665
	/organism="Onchocerca"

```
/strain="Kumba, Cameroons"  
/db_xref="taxon:6282"
```

```
/clone_lib="Onchocerca volvulus molting L3 larva cDNA  
(SL96MLW-OvmL3)"  
/dev_stage="molting L3"
```

```

/note="Vector: lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies

```

in IMDM+ NCTC 135 and collected after day 1, 2, or 3 culture. L3 of *O. volvulus* molt to fourth-stage larvae day 5 in culture. mRNA was isolated from approximately 1000–1500 larvae from each day 1–3

In culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase and DNA pol I. The library was constructed in the  $\lambda$  BamHI-Xba vector and had  $1 \times 10^6$  independent

The library is available from Dr. Sara Lustigman (email: [slustigman@post.queensu.ca](mailto:slustigman@post.queensu.ca)).

BASE COUNT	186 a	129 c	136 g	213 t	1 others
ORIGIN					

Query Match	22.0%	Score 250.12	DB 2	Length 100%
Best Local Similarity	84.1%	Pred. No. 5.8e-98		
Matches 451; Conservative	0;	Mismatches 84;	Indels 1;	Gaps

117 ATGCTATTCGGGTGATAACGGTGTAGAAGTGAAACGAAATTGATGTGGTCCACAT 176

Db 177 CAATTACTGTCACCTTAACTACTCGTAATCCATTGGAAGACATGTATACGTGAAAAGGCT 236

Db 237 TATACGATCAGCGCAGGATGCCGAATGATGAGAGGTGGACCTCAGGTAGCCGGAATTGCAAC 296

Db 297 TTCCGTTTGATTCCTGATATGACAGCTACAGCTTCGTAATCCAGCTGATTTTG 356

Db 357 TCACATCAGTTGTTGCATCATTCATCCATGTTTGGACAAAGTTGATCGACAT 416

Db 417 ATCGATACATGCTTTACATGGAGCGTGATNAGACAGTTAGCGCTCAACCTTGAAGCTTTT 476

Db 477 CCGAATGACAACCTGCATTTGCCAACACAATTTGCCGATGCCGTATGCGGATATGAGA 536

Db 537 TTCCTGATGGTGGACCAATCCGGACAACCTGTTCAATTGCTACCAATTGGTCAACCACTGT 596

QY 530 ATCATATAATGACATGGATTCGAAACCGTTACTTCTGCGCGGTTTCAT 585  
 Db 597 -CCATAATGACATGTGATTCGAAACAGTTGATACATTCCTGGCTGTGGACAT 651

RESULT 7  
 BF727527 740 bp mRNA linear EST 08-JAN-2001  
 LOCUS BF727527  
 DEFINITION SMOV3MCA50E11SK Onchocerca volvulus molting L3 larva cDNA (SL96MLM-Ovml3)  
 ACCESSION BF727527  
 VERSION BF727527  
 KEYWORDS EST.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filariidae; Onchocercidae; Onchocerca.  
 REFERENCE 1 (bases 1 to 740)  
 AUTHORS Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
 TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: Bluescript SK.  
 FEATURES  
 source location/Qualifiers  
 1..740  
 /organism="Onchocerca volvulus"  
 /strain="Kumba, Cameroons"  
 /db\_xref="taxon:6282"  
 /clone="SMOV3MCA50E11"  
 /clone\_1lb="Onchocerca volvulus molting L3 larva cDNA (SL96MLM-Ovml3)"  
 /dev\_stage="molting L3"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda Uni-ZAP XR; Site.1: Eco RI; Site.2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+NCFC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda uni-zap XR vector and has 1 x 10E6 independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustig@nyc.org)."

BASE COUNT 217 a 140 c 157 g 226 t

ORIGIN

Query Match 33.4%; Score 387.2; DB 12; Length 740;  
 Best Local Similarity 82.0%; Pred. No. 41e-97;  
 Matches 482; Conservative 0; Mismatches 103; Indels 3; Gaps 3;

QY 50 ATTCATTCGCGTTCGATGTCGAGTCGACCAATTCGAATTCGACCACTT 109  
 Db 147 ATGCATTCGCGTTCGATGTCGAGTCGACCAATTCGAATTCGATTCCTCACT 206

QY 110 CATATACATCAATTTAATACAGTATGATTCGAGAGCATGTTATGTGAAGGTC 169  
 Db 207 CATATACATCAATTTAATACAGTATGATTCGAGAGCATGATATCGTAAGGCT 266

QY 170 TTATGATCAAGAAGTTGCCGTAATGANGAAGTGAGCATGTAAGTTCCCGAATTCAC 229  
 Db 267 TATACATGACGAGATGCCGAATGATGAGTGCAGCTAGTAGCCGGAATTCAGC 326

QY 230 TTCCATTTGATTCATGCAATGTTGCCGTCACAGATCTCGAATCCACGGTATTTTGG 289  
 Db 327 TTCCGTTGATTCCTGTAAGTAGACACAGTCGTTAAATCCACGGTATTTTGG 386

QY 290 TAACCAACATGTTGTCATATTCGTTTCATCCATTTATTTGTTACCAAGTTGATGTCAT 349  
 Db 387 TCATCATCAGTTGTGTCATCTCATTCATCCATTCATGTTTGTGACAAAAGTTGATGACAT 446

QY 350 ATGAGTACATGCTTTTACATGANGAGCTGATTAACAGTATGATGACATGAGTAT 409  
 Db 447 ATGCAATACATGCTTTTACATGANGAGCTGATTAACAGTATGATGACATGAGTAT 506

QY 410 CTGAATTCACATGCTTTTCAACATCAATTCCTCCGATCCGATGATCCGTTANGAAN 469  
 Db 507 CCGAAATGACATGCTATTTGCAACACAAATGTTACCGATGCTGATCCGATATGAGA 566

QY 470 TTTTGGATGTCGACCAACCGTCGTCACAGTTCATTTGCTATTCATGTCACAGCTTT 529  
 Db 567 TTCTTATGTCGACCATCCGAC -ACGTTTCATTTGCTACATTTGTCACACAGTG 625

QY 530 ATCATATAATGACATGGATTCGAAACCGTTGATCTTTCGCGGTTTCATTCCT 589  
 Db 626 TCATATAATGACATGTGATTCG -AACAGTTGATTCATCTTG -GGGCTGTGACATTCAT 683

QY 590 GCTTGTGATGATGTAACGGTGATGATGTAATTCGTAATTCGTC 637  
 Db 684 GCTTGTGATGATGTAACGGTGATGATGTAATTCGTAATTCGTC 731

RESULT 8  
 BF400424 648 bp mRNA linear EST 28-NOV-2000  
 LOCUS BF400424  
 DEFINITION SMOV3MCA50E11SK Onchocerca volvulus molting L3 larva cDNA (SL96MLM-Ovml3)  
 ACCESSION BF400424  
 VERSION BF400424  
 KEYWORDS EST.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filariidae; Onchocercidae; Onchocerca.  
 REFERENCE 1 (bases 1 to 648)  
 AUTHORS Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
 TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: Bluescript SK.  
 FEATURES  
 source location/Qualifiers  
 1..648  
 /organism="Onchocerca volvulus"  
 /strain="Kumba, Cameroons"  
 /db\_xref="taxon:6282"  
 /clone="SMOV3MCA45G11"  
 /clone\_1lb="Onchocerca volvulus molting L3 larva cDNA (SL96MLM-Ovml3)"  
 /dev\_stage="molting L3"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda Uni-ZAP XR; Site.1: Eco RI; Site.2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS



EST.  
pig roundworm.  
Ascaris suum  
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea  
Ascarididae; Ascaris.  
1 (bases 1 to 646)  
McCarte, J., Clifton, S., Chapell, B., Pape, D., Martin, J., Wylie, T.,  
Danton, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,  
Gibbs, N., Ritter, E., Bennett, J., Franklin, C., Tesagrelshvili, R.,  
Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steplee,  
M., Allen, M., Person, S., Sweller, T., Harey, N., Schurr, R., Kohn, S.,  
Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and  
Johnson, P.

**TITLE** The Washington Univ. Nematode EST Project, 1999  
**JOURNAL** Unpublished (1999)

The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 285 1800

Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1. Seq primer: -40RP from Gibco  
High quality sequence, stop: 482.

```

source
1. 646
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_id="Ascaris suum I4 pSPORT1 Zarlena v1"
/dev_stage="L4"
/lab_host="DH5-alpha"
/notes="Vector: pSPORT1 (Life Technologies); Site.1: NotI;
Site.2: SalI; The library was supplied by Dr. Dante
Zarlena of the USDA Immunology and Disease Resistance Lab
in Beltsville, MD. The cDNAs were made from 21 day I4
proximal (isolated from jejunum) and directionally cloned
into the NotI/SalI sites of pSPORT1."
BASE COUNT
152 a 168 g 152 t 1 others
      173 c

```

Query Match	32.5%	Score	376.8	DB	13	Length	646
Best Local Similarity	74.4%	Pred. NO.	3.2e-94				
Matches 474; Conservative	0	Mismatches	163	Indels	0	Gaps	0

QY 32 CACATTAATGCATTCTCTTATTGCAATTCGGGTGACAAATGCTGTGGAAGGTGACCAGAAA 91  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 6 CTCTTTTTCGGGCTCGCCGACACTATTCTCTTGCAACAAGGAGTGTGMAAGAAGGCCGGAGAGA 65  
  
QY 92 TTGTAATGAGCAACTTCATAAATCATTTTTTAATACAGTAATGCATTTCGAAAGAC 151  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 66 TTGTAATGTGGTCTCAACGTCATCAGGGCACTTCAACACAGAAATCATTCGAGGGTCT 125  
  
QY 152 ATGTTTATGTTAAAGGCTTTATGATCAAGAAGGTGGCGTAAATGATGAAGGTGAAGCTC 211  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 126 ATGATATCGTBAAGGACCTTATATANTCAAAGGGATGCGCACGACGACGAGAGGTGGACGTC 185

[illegible]

Db	246	ATCCCTGATGATCTTTTGTCAGCAGCAGCGTTGTCATCTCTTCATCGGCTCTTCATCA	305
Db	332	CCAAAGTGTGATCGGCATATCGATAGTACATGCTTTTACATGGAAGCTATATAACAGTTA	391
Oy	306	CCAAAGTGTGACCGGTGCGCTACCGCTCCAAATGCTTCTACATGAGAGCCGACCAAGACTGTCA	365
Db	392	GTCGACACGATTTGAGGTATCTGTAATTCACAACTGGTTTTCAAGCTCCAAATTTGCTCCGATGC	451



Db 366 GTACGCAAACTGAGTGTGAGATTGACACCGCTTCCAGCAATGCTCTATGC 425  
QY 452 CAGTATGCCCTTATGAAATTTGGATGGACCAACCGGTCAACAGTTCATTGCTA 511  
Db 426 CAGTCTGCAATAGAGATTCTGACGAGAGTCCACAGGACGCCATACAGTCCAA 485  
QY 512 TCATTGCTCAGCCAGTTTATCATTAATGACATGCGATTCTGAAACCGTTGATCTTCT 571  
Db 486 CCATCGGTCAGCAATATACACAGGAGCTTGCATTCGGAACCTGTGATACGTTCT 545  
QY 572 GCCCGGTTGCCATTCCTGTTTGGATGATGTAACGGTGTACTGTAATTCCTA 631  
Db 546 GTGCTGTCTTCACTACTGCTGTGATGATGACAGCGGTGATGATCAGATCCTTC 605  
QY 632 ATGCTGATGATGCTGCTTGTATTAATATTGCTTAA 668  
Db 606 ACGAGGAGGGTGTGCGCTTGACAAATACCTGCTGAA 642

RESULT 10  
Bg809012 747 bp mRNA linear EST 22-MAY-2001  
LOCUS SMOV3MCA1144SK Onchocerca volvulus molting L3 larva cDNA  
DEFINITION (SL96MLM-OvML3) Onchocerca volvulus cDNA clone SMOV3MCA1144 5',  
mRNA sequence.  
ACCESSION Bg809012 GI:14180004  
VERSION Bg809012.1 GI:14180004  
KEYWORDS EST.  
SOURCE Onchocerca volvulus.  
ORGANISM Onchocerca volvulus.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Onchocerca.  
REFERENCE 1 (bases 1 to 747)  
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.  
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
JOURNAL Unpublished (1997)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genomesmith.edu  
Seq primer: pluscript SK.  
Location/Qualifiers  
FEATURES  
source 1..747  
/organism="Onchocerca volvulus"  
/strain="Kumba, Camerouns"  
/db\_xref="taxon:6282"  
/clone="SMOV3MCA1144SK"  
/clone\_1ib="Onchocerca volvulus molting L3 larva cDNA (SL96MLM-OvML3)"  
/dev\_stage="molting L3"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: Lambda Uni-ZAP XR; Site.1: Eco RI; Site.2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ML3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10<sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigmen@nc.org)."

BASE COUNT 210 a 137 c 159 g 241 t  
ORIGIN  
Query Match 32.4%; Score 376.2; DB 12; Length 747;  
Best Local Similarity 80.4%; Pred No. 4.9e-94;  
Matches 514; Conservative 0; Mismatches 118; Indels 7; Gaps 6;  
QY 50 ATTGATTCGCGGTGCAATGGTGTGCAAGGTGAGCCAGAAATGATGTGACCACTT 109  
Db 109 ATGCTATTCGCGGTGTAATACGGGTGTAGTGAACACAGAAATGATGTGCAACAT 168  
QY 110 CAATTAACATCAATTTTAATACAGTAATGATTCAGAGACATGTTTATGAAAGTTC 169  
Db 169 CAATTAACATCAATTTTAATACAGTAATGATTCAGAGACATGTTTATGAAAGTTC 228  
QY 170 TTATGATGCAAGAGGTGCGCGTAATGATGATGAGGTGAGGTGAGGTGAGGTGAG 229  
Db 229 TATGATGCAAGAGGTGCGCGTAATGATGATGAGGTGAGGTGAGGTGAGGTGAG 288  
QY 230 TTGCAATTCATTCATGCAATGTTGCGGTGACACAGATCTGTGATCCAGTATTTTG 289  
Db 289 TTGCAATTCATTCATGCAATGTTGCGGTGACACAGATCTGTGATCCAGTATTTTG 348  
QY 290 TACAGCAACGTGTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 349  
Db 349 TACAGCAACGTGTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 408  
QY 350 ATGAGTACAAATGCTTTTACATGATGATGATGATGATGATGATGATGATGATGAT 409  
Db 409 ATGAGTACAAATGCTTTTACATGATGATGATGATGATGATGATGATGATGATGAT 468  
QY 410 -CTGAATACAACTGC-CTTTCAAACTCAATTTGCTCCGATGCCAGTATGCCGTTATGA 467  
Db 469 CCGGAATGACAACTGCCTTTTCAAACTCAATTTGCTCCGATGCCAGTATGCCGTTATGA 528  
QY 468 AATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 527  
Db 529 GATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
QY 528 TTATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 586  
Db 589 TGTCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648  
QY 587 CCTGCTTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644  
Db 649 CATGCTTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
QY 645 TGCTCTT--GATTAATATTTGCTTAATTAATTTGGAATAT 681  
Db 709 GGGCTTTGGACAAATTTTGTCTCAACACCTTGATAT 747

RESULT 11  
AA701731 628 bp mRNA linear EST 19-DEC-1997  
LOCUS SMOV3MCA1144SK Onchocerca volvulus molting L3 larva cDNA  
DEFINITION (SL96MLM-OvML3) Onchocerca volvulus cDNA clone SMOV3MCA1144 5',  
mRNA sequence.  
ACCESSION AA701731  
VERSION AA701731.1 GI:2704931  
KEYWORDS EST.  
SOURCE Onchocerca volvulus.  
ORGANISM Onchocerca volvulus.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Onchocerca.  
REFERENCE 1 (bases 1 to 628)  
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.  
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
JOURNAL Unpublished (1997)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith



Db	148	TATAGATAGAGCAGGATGCCGAATGATGAAGGTGAGGCTGACGTAGCGGAAATTGAAC	207
Qy	230	TTTCATTTATTCATGATGCAATGTTGCCCGTACAGATCTCTGAATCCAGCGGTATTTTG	289
Db	208	TT	267
Qy	290	TAAACAACACTGTGTGTCATTCGTTTCATCCATCTATTTTGTACCAAGTGTGATCGTCAAT	349
Db	268	TCACATCAAGTGTGTGTGTCATCTCATTCATTCATCTGTTTGTGACAAAAGTTGATCGACAT	327
Qy	350	ATCGATGACAAATGCTTTTACATGGAAGCTGATAAAACAGTTAGTGCACAGATTGAGGAT	409
Db	328	ATCGAAT-CAATGCTTTTACATGGAAGCTGATAAAGACAGTTAGCGCTCAACTTGAAGCTT	386
Qy	410	CTGAATATGCAACTGCTTTTTCAAACATCAATTTGCCCATGCGCAGTATCCGCTTATGAAA	469
Db	387	CCGAATATGCAACTGCTTTT-GCACACAAATTTGTACCGATGCTGTATCCGATATGAGA	445
Qy	470	TTTTGATGATGATGACCAACCGGTCAACAGTTCATTAATTTCTATTCATTTGTCACCGAGTT	529
Db	446	TTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	504
Qy	530	ATCATTAATGACATGCGATTCCTGGAACCGTGTGATCTTTCGCGCGGTTGTCATTCCT	589
Db	505	ACCATAAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	561
Qy	590	GCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	649
Db	562	GCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	618
Qy	650	TTGATTAATAT 660	
Db	619	TGGACATTTT 629	
RESULT 13			
LOCUS	BF599177		
DEFINITION	BF599177 774 bp mRNA linear EST 13-DEC-2000		
	SMOV3MCAM49G08SK Onchocerca volvulus molting L3 larva cDNA		
	(SIJ96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCAM49G08 5',		
	mRNA sequence.		
ACCESSION	BF599177		
VERSION	BF599177.1 GI:11695516		
KEYWORDS	EST.		
SOURCE	Onchocerca volvulus.		
ORGANISM	Onchocerca volvulus		
	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;		
	Onchocerca volvulus.		
REFERENCE	1 (bases 1 to 774)		
AUTHORS	Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.		
TITLE	Genes expressed in molting L3 larvae of Onchocerca volvulus		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Steven A. Williams		
	Molecular Parasitology		
	Smith College Department of Biological Sciences		
	Department of Biological Sciences, Clark Science Center, Smith		
	College, Northampton, MA, 01063, USA		
	Tel: 4135853826		
	Fax: 4135853786		
	Email: genome@smith.edu		
FEATURES	Seq primer: pBluescript SK.		
SOURCE	Location/Qualifiers		
	1..774		
	/organism="Onchocerca volvulus"		
	/strain="Kumba, Camerouns"		
	/db_xref="taxon:6282"		
	/clone="SMOV3MCAM49G08"		
	/clone_id="Onchocerca volvulus molting L3 larva cDNA		
	(SIJ96MLM-Ovml3)"		
	/dev_stage="molting L3"		
	/lab_host="XLI-Blue MRF"		
	/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:		
	Xho I; Filarial nematode parasite of humans. Third-stage		

BASE COUNT	233 a	149 c	159 g	233 t
ORIGIN	<p>larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+NCIC 135 and collected after day 1, 2, or 3 in culture. L3 of <i>O. volivulus</i> molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligod(T) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10<sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigman@dc.org)."</p>			
Query Match	31.9%	Score 369.8	DB 12	Length 774
Best Local Similarity	81.2%	Pred. No. 3.1e-92		
Matches	491: Conservative	0: Mismatches	107: Indels	7: Gaps
OY	50	ATTCGATTCGGGTGACATGGTGTGCAAGGAGACCAAAATTTGAAGCACCACT	109	
DB	170	ATGCAATTCGGGTGATAACGGTGTGAAGAGGACCAAAATTTGAATGGTCCACACT	229	
OY	110	CAATAACAATCAATTTAATACAGCTAATGATCGAAGACATGTTATGTGAAGGTC	169	
DB	230	CATTACTGTCACATTTAATCTGTAATCCATTTGAAGACATGATACGTGAAGGCT	289	
OY	170	TTTATGATTAACAAGGTCCGTAAATGAAAGGTGACGTAAAGTTGCCGAATTTAC	229	
DB	290	TATAGATTCAGGACAGATGCGCAAAATGATGAAGGTGACGTACGGCAATTTGAC	349	
OY	230	TTTCAATTTGATTCATGATGTTGGCGGACACGATCTGTAATCAGACGTGATTTTG	289	
DB	350	TTTCGTTGATTTCCGTAAATGACACGTCACGTTCCGTTAAATCCAGGTGATTTTG	409	
OY	290	TAAACAACAATGTTGTCAATTTGTTTCATCCATTAATTTGTTACCAAGTTGATCGTAC	349	
DB	410	TCACATCAATGTTGTTCATCTCATCTTCATCCATTCATGTTGTGCAAAAGTTGATCGACAT	469	
OY	350	ATTCGATTCATGCTTTTACATGGAAGCTGATTAACAAGTGTGTCACAGATTTAGGTAT	409	
DB	470	ATCGAATCAAAAGCTTTTACATGGAAGCTGATTAACAAGTGTGTCACAGTGTG	529	
OY	410	-GTGAATCACAAAGCTTTTCAAACTCAAAATTTGCCGATCCGATTAATGCAATGAA	468	
DB	530	CCCGAAATGACACAGCTGATTTGCAACCAAAATTTG-CCGATGCCGATTAATGCGATATGAG	588	
OY	469	ATTTTGATGATGGAGCAACCGGTCAACC-AGTTCAATTTGCTATCATTTGTCACCCAGT	527	
DB	589	ATTCTTGATGGTGGACCAATCCGGGACCACTTTTCAATTTGCTACATTTGGTCAACAGT	648	
OY	528	TTATCATTAATGACATCGATCTGGAACCGTTG--ATACTTCTGCGCGGTGTCCAT	585	
DB	649	GGTCACTAAATGACATGATGATTTGGAACAGTTGATACCATCTGCTGCTGATCAT	708	
OY	586	TCCTGCTTTGTCGATGATG--TAAAGGTGATCTGGAATTTCTAATGCTATGAT	643	
DB	709	TCATCTTCTGTTGATGATGATGATTAACCGGTGATTAAGGGGAAATTAATTTAAATGCTTGAT	768	
OY	644	GTGCT 648		
DB	769	GGGT 773		
RESULT 14	AA618895 672 bp mRNA linear EST 12-NOV-1997			
LOCUS	AA618895			
DEFINITION	SMOVB3C4A1879SK Onchocerca volivulus molting L3 larva cDNA (S1956LM-Ovml3) Onchocerca volivulus cDNA clone Sml13C01879 5', mRNA sequence.			

ACCESSION	AA618895
VERSION	AA618895.1 GI:2522771
KEYWORDS	EST.
SOURCE	Onchocerca volvulus.
ORGANISM	Onchocerca volvulus Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
REFERENCE	Onchoceroidae. Onchocerca. 1 (bases 1 to 672)
AUTHORS	Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S. Genes expressed in molting L3 larvae of Onchocerca volvulus Unpublished (1997)
TITLE	Contact: Steven A. Williams
JOURNAL	Molecular Parasitology Smith College Department of Biological Sciences Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA Tel.: 4135853826 Fax: 4135853786 Email: genome@smith.edu Seq primer: pBluescript SK. Location/Qualifiers 1..672
FEATURES	/organism="Onchocerca volvulus" /strain="Kumba, Cameroons" /db_xref="taxon:6282" /clone="SmmL3C01879" /clone_1fb="Onchocerca volvulus molting L3 larva cDNA (SL96MLM-OvmL3)" /dev_stage="molting L3" /lab_host="XLI-Blue MRP" /note=Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dt) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10^6 independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams The library is available from Dr. Sara Lustigman (email: slustigm@cyc.org).
BASE COUNT	198 a 132 c 129 g 202 t 11 others
ORIGIN	
Query Match	31.4%; Score 364.2; DB 9; Length 672; Best Local Similarity 81.1%; Pred. No. 1,le-90;
Matches	417; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
OY	50 ATTCGATTCCGGTGACAATGGTGTGCAGAGGTGACCAGAATAATGATGGACAATT 109
Dy	159 ATCGTAATCCGGTGATGAACGGTAGAAGGAGCAACAAATTTGATGTGCCAACAT 218
OY	110 CAATATCATCAATTTTTAATACACTAATGCAATTTGGAGAGCAATGTTTATGTGAAGATC 169
Dy	219 CAAATTAAGTCAACTTATACTGTATACCTTAATCCCCATTTTGAAGACATGTATACGTGAAGGGCT 278
OY	170 TTTATGATCAAGAAGGTTGCCGTATGATGAAGGTGAGCGTCAAGTGGCCGGAATTTCAC 229
Dy	279 TATAGATCACGAGCAAGATGCCAATATGATGAAGGTGAGCGTGAAGTACGCCGGAATTGAAC 338
OY	230 TTTCATTTGATTCATGACAATGTGGCCGTACACGATCTCTGCAATCCACGTGTATTATTGG 289
Dy	339 TTCGTTGATTCCTGTATATGAGACGTAACGTTCCGTTAATCATCAGCTGTGATTTTGG 398
OY	290 TAACACAACATGTGTGCAATTTGTTTCATCCAATTTATTTTTCACAAAGTTGATCGCAT 349
Dy	399 TCACATCAGTGGTTTCNNCTATTCATCCATCTGTTGTCGACAAAAGTTGATCGACAT 458

QY	350	ATTGAGTGCATTCGTTTACATGSAAGCGMTAAACAGTTAGTGCACAGATTGAGGT	409
Db	459	ATCGAATCAAAATGCTTTTACATGSAAGCTGATTAACAGTTAGTGCACATTGAATTT	518
QY	410	CTGAATTCACAACTGCTTTTCAAACTCAAAATGTGCCATGCGCAGTATGCCCTTATGAA	469
Db	519	CGGAATTCACAACTGCTTTTTCGTCGAACTGATGATGATGATGATGATGATGATGATG	578
QY	470	TTTGGATGTTGGAGCAACCGGTCAACCAAGTTCAATTTGCTATCATTTGGTCAGCAGTTT	529
Db	579	TTCTTGATGTTGGAGCAACCGGTCAACCAAGTTCAATTTGCTATCATTTGGTCAGCAGTTT	638
QY	530	ATCATTAATGACATGCGATTTCTGAACCGTTGA	563
Db	639	TCCATTAATGACATGCGATTTCTGAACCGTTGA	672
RESULT 15			
LOCUS	BM284858	612 bp	mRNA linear EST 28-DEC-2001
DEFINITION	km9ph05.y1 Ascaris suum L4 pSPOR1 Zarlena vi Ascaris suum cDNA 5'		
	similar to TR:Q93115 Q93115 CUT-1-LIKE CUTICLIN PROTEIN PRECURSOR.		
ACCESSION	BM284858		
VERSION	BM284858.1	GI:17993900	
KEYWORDS	EST.		
SOURCE	plg roundworm.		
ORGANISM	Ascaris suum		
	Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea		
REFERENCE	1 (bases 1 to 612)		
AUTHORS	McCartner,J., Clifton,S., Chippelli,B., Page,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,T., Gibbons,M., Rilter,E., Bennett,J., Franklin,C., Tsagarisvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swalter,T., Harvey,N., Schurk,R., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.		
TITLE	The Washington Univ. Nematode EST Project, 1999		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: McCartner JP The Washington Univ. Nematode EST Project, 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu The library was suppld by Dr. Dante Zarlena of the USDA Immunology and Disease Resistance lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPOR1. Seq primer: -40RP from Glpbo High quality sequence stop: 480. Location/Qualifiers 1..612 /organism="Ascaris suum" /db_xref="taxon:6253" /clone_lib="Ascaris suum L4 pSPOR1 Zarlena v1" /dev_stage="L4" /lab_host="DH5-alpha" /note="Vector: pSPOR1 (Life Technologies); Site_1: NotI; Site_2: SalI; The library was supplied by Dr. Dante Zarlena of the USDA Immunology and Disease Resistance Labh in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPOR1."		
FEATURES			
source			
BASE COUNT	165 a	140 c	147 g 160 t
ORIGIN			
Query Match	31.3%;	Score 363.6;	DB 13; Length 612;
Best Local Similarity	75.8%;	Prod. No. 1.5e-90;	
Matches 450:	Conservative 0;	Mismatches 144;	Indels 0; Gaps 0;

```
QY 41 CATTGCTATTTGGATTCGGTTCACATGATGTCAGAGTGCAGCAAAATGATGTG 100
    ||||| || ||||| ||||| || ||||| || ||||| || ||||| |||||
Db 15 CATTACTGTGGGATCCCGGTTGATAACGGCTTGAGGAGAAACCGAAATGAAATGTG 74
QY 101 GACCACTTCATACATCAATTTTAAATACAGTATGATGATGCAAGCATTTTATG 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 GCGCCACCTCAATACAGTCACTTTTACCTGAAATCAATTTGAGGGACGCTTACG 134
QY 161 TGAAGGCTTTATGATCAAGAGGTTGCCGTAATGATGAGAGTGCAGTTCAGTTC 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 TAAAGAGCTCTATGATCAAGAGGATGCCGTAACGATGATGAGAGTTCAGTTC 194
QY 221 GAATTCACCTTCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 GATTCACGTTACCGTTCGATGATGATGATGATGATGATGATGATGATGATGAT 254
QY 281 GATTTTGTAAACACACAGCTGTTGATTCATTCATTCATTCATTCATTCATTC 340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GCATCTTCGTGACGACGATGCTGTGATCTCATTTTCACTCTTCGTCACCAAGT 314
QY 341 ATCGTCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 ATCGACCTATCGCATTCAGTGTCTTACATGAGAACCGACAGACATCAGTACAG 374
QY 401 TTGAGGATCTGAATACACACTGCTTTTCAAACTCAAAATGTCGCGATGCCAG 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 TGAAGTGTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 434
QY 461 GTTATGAAATTTTGGATGATGATGATGATGATGATGATGATGATGATGATG 520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 GATACGAGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 494
QY 521 AGCCAGTTATCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 AACAGATATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
QY 581 TCCATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 TCCATTCATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 608

RESULT 16
A1322117 639 bp mRNA linear EST 22-DEC-1998
LOCUS A1322117
DEFINITION SMOV3MCA012G08SK Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA012G08 5',
mRNA sequence.
ACCESSION A1322117
VERSION A1322117.1 GI:4056268
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 639)
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL Unpublished (1997)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
FEATURES
Source
1..639
Location/Qualifiers
/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
```

```
/clone="SMOV3MCA012G08"
/clone_lib="Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3)"
/dev_stage="molting L3"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda Uni-Zap XR; Site.1: Eco RI; Site.2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroon (forest strain). The L3 were cultured in 20% FCS
in IMDM+NCM 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNA pol I. The library was constructed in the lambda
uni-zap XR vector and has 1 x 10E6 independent
recombinants and the average insert size is ~1200 bp. The
library was constructed by Sara Lustigman and Michelle
Lizotte-Waniewski in the laboratory of Dr. S. A. Williams.
The library is available from Dr. Sara Lustigman (email:
slustig@nbc.org)."
```

BASE COUNT 197 a 130 c 121 g 191 t

ORIGIN

Query Match 30.9%; Score 358.6; DB 9; Length 639;  
Best Local Similarity 84.5%; Pred. No. 3.8e-89;  
Matches 403; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```
QY 50 ATTGATTCGGTGTGACAAATGATGATGATGATGATGATGATGATGATGATGAT 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 ATGCTATTCGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 221
QY 110 CATTAACATCAATTTTATACAGTAATGATGATGATGATGATGATGATGATGAT 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 CATTAACATCAATTTTATACAGTAATGATGATGATGATGATGATGATGATGAT 281
QY 170 TTTATGATCAAGAGTGGTTCGTAATGATGATGATGATGATGATGATGATGATG 229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 282 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 341
QY 230 TTCCATTTGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 TTCCGTTGATTCCTGATATGATGATGATGATGATGATGATGATGATGATGAT 401
QY 290 TACCAACAACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 TCCATCACTGTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 461
QY 350 ATGAGTACATGCTTTTACATGATGATGATGATGATGATGATGATGATGATGAT 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 ATGATACAAATGCTTTTACATGATGATGATGATGATGATGATGATGATGATG 521
QY 410 CTGAATACAACTGCTTTTACAACTGATGATGATGATGATGATGATGATGATG 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 CCGAATGACAACTGCTTTTACAACTGATGATGATGATGATGATGATGATGATG 581
QY 470 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 582 TTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638

RESULT 17
AA668071 610 bp mRNA linear EST 20-NOV-1997
LOCUS AA668071
DEFINITION SMOV3MCA020C01SK Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA020C01 5',
mRNA sequence.
ACCESSION AA668071
VERSION AA668071.1 GI:2629570
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus
```









ACCESSION		similar to TF:Q93115 Q93115 CUT-1-LIKE CUTICLIN PROTEIN PRECURSOR ; mRNA sequence.
VERSION		BM284979
KEYWORDS		BM284979.1 GI:17994021
SOURCE		EST.
ORGANISM		pig roundworm. Ascaris suum Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
REFERENCE		1 (bases 1 to 624) McCarteer,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucab,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvill,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE		The Washington Univ. Nematode EST Project, 1999
JOURNAL		Unpublished (1999)
COMMENT		Contact: McCarteer JP The Washington Univ. Nematode EST Project, 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1. Seq primer: -40RP from Gibco High quality sequence stop: 483.
FEATURES		Location/Qualifiers 1..624 /organism="Ascaris suum" /db_xref="taxon:6253" /clone_lib="Ascaris suum L4 pSPORT1 Zarlenga v1" /dex_stage="L4" /lab_host="DH5-alpha" /note="Vector: pSPORT1 (Life Technologies); Site_1: NotI; Site_2: SalI; The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1."
BASE COUNT	156 a 155 c 167 g 146 t	
ORIGIN		
Query Match	28.8%; Score 333.8; DB 13; Length 624;	
Best Local Similarity	71.3%; Pred. No.5.2e-82;	
Matches 440:	Conservative 0; Mismatches 177; Indels 0; Gaps 0;	
Db		
32	CACATTAATGCATTGCTTATTCGATTCGGGTGACAAATGTTGCGAGGTGACCCAGAAA	91
7	GCGTCGCGGATTCGCTTCTGCGATTCCAAGTCGAACAATGGAGTCCAGGAGAGGCCGAGA	66
OY	92 TTGAATGTGACCACTTCAATTAACAAATCAATTTTAATAACAGTAATGATTCGAGAAC	151
Db	67 TCGAGTGCGTCGCCCTTCAATCACAGTCAACTCAACACTCGAACCATTCCAAAGGCG	126
OY	152 ATGTTTANGTGAAGAGCTTTATATGATCAGAAGAGTTCGCCGTATATGTAAGGTGACGTC	211
Db	127 ACGTTTAGTGAAGGATGTGACGTACGTCAGGAAGATGTCGAAACGACMAAACGGCGCTA	186
OY	212 AAGTTCGCGGAATTCATCTCATCTTGAATTCAGCAATGTTCCGGTACACAGTCMGTA	271
Db	187 GTGTGCGGGAATTAACGCTGCCGTTCCGACAGTTGTAATGTGACACGAGACTCGTTGTTGA	246
OY	272 ATCCAGCTGTGATTTTTGTAACAACAACACTGTTGTCAATTCGTTTCATTCATTAATTTGTTA	331
Db	247 ATCCAGAGGCGTCTTCGTAAGTACGAGGTGTCATCATTCATTCATCCCTTTTCGTTA	306
OY	332 CCAAAGTGATTCGTCATATCGAGTACATGCTTTTACATGGAAGACTGATAAACAAGTTA	391

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
D0	307	CCAAAGTACACAGACACATACCGCATTCAGTGTCTTCTACATGTGAGGCCGATTAAGACCGCA	366								
QY	392	GTGCACAGATTGAGGTATCTGAAATCACACACTGCTTTTCAAATCTCAATTTGCCGATGC	451								
D0	367	GCACACTACTGGAAGTGTCCGAAATACACACCGCTTCCAAACGAGGTGTTCGATGC	426								
QY	452	CAGTATGCCCTTATGAAATTTTGGATGTGGACCAACCGGTTCACACACTTAAATTTGGTA	511								
D0	427	CCATCTGCAGATATGAAATTAACCTTGACGGTGGCCACACAGACAGCTTATCCAAATTTGGCA	486								
QY	512	TCATTGTGCACACAGTTATCATTAATGAGATCGATTCGAAACCGTTATCTTCT	571								
D0	487	CTATTGGACACAGGTATATACACAAATGGAATTTGGATTCGAGACTGTTACACCTTCT	546								
QY	572	GCGGGTGTGTCATTCTCTGCTTTGTGATGATGTTACGATGATATGTGAAATTTCTTA	631								
D0	547	GTGACAGTGTGCACACTCATCTCTGTTGACGACGCAATGGCGATACAGTGCAGCTTGA	606								
QY	632	ATGCTGATGATGCTTCT 648									
D0	607	ACGAGGAGGATGGCT 623									
RESULT 21	BP482099										
LOCUS	BP482099										
DEFINITION	BP482099	554 bp	mRNA	linear	EST 06-DEC-2000						
ACCESSION	SMOV3MCAM47608SK	Onchocerca volvulus molting L3 larva cDNA									
VERSION	(SL96MLM-Ovml3)	Onchocerca volvulus cDNA clone SMOV3MCAM47608 5'									
KEYWORDS	BP482099	mRNA sequence.									
ORGANISM	BP482099.1	GI:11565400									
SOURCE		Onchocerca volvulus.									
ORGANISM		Onchocerca volvulus									
REFERENCE		Eukaryotes: Metazoa: Nematoda: Chromadorea: Spiruridae: Filarioidea:									
AUTHORS		Onchocercidae; Onchocerca.									
TITLE		1 (bases 1 to 554)									
JOURNAL		Williams, S.A., Lizotte-Waniewski, M., Taney, S. and Lustigman, S.									
COMMENT		Genes expressed in molting L3 larvae of Onchocerca volvulus									
		Unpublished (1997)									
		Contact: Steven A. Williams									
		Molecular Parasitology									
		Smith College Department of Biological Sciences									
		Department of Biological Sciences, Clark Science Center, Smith									
		College, Northampton, MA, 01063, USA									
		Tel: 4135853826									
		Fax: 4135853786									
		Email: genome@smith.edu									
		Seq primer: pBluescript SK.									
		Location/Qualifiers									
FEATURES		1..554									
SOURCE		/organism="Onchocerca volvulus"									
		/strain="Kumba, Cameroons"									
		/db_xref="taxon:6282"									
		/clone="SMOV3MCAM47608"									
		/clone_1bp="Onchocerca volvulus molting L3 larva cDNA									
		(SL96MLM-Ovml3)"									
		/dev_stage="molting L3"									
		/lab_host="XLI-Blue MRP"									
		/note="Vector: lambda uni-ZAP XR; Site_1: Eco RI; Site_2:									



Query Match	26.9%	Score 312.4	DB 9	Length 537
Best Local Similarity	84.2%	Pred. No. 2.9e-76		
Matches 352	Conservative	0	Mismatches 66	Indels 0
			Gaps 0	
Qy 50	ATTCGATTCGCCGTGACATGCTGTGCGAAGGTGACCCAGAAATGTGATGGACCACTT	109		
Db 120	ATGCAATTCGGTTGATATACGGTGTAGAAAGGTGAACCGAAATTTAAATGGTCCACACT	179		
Qy 110	CAATPAACAATCAATTTTAATACACGTAATGCAATTCGAAAGACATGTTATATGAAAGTGC	169		
Db 180	CAATTAATCTCAACTTATATCTGTAATCCATTTTGAAGGACATGTATACGTGAAGGT	239		
Qy 170	TTTATGATCAAGAGGTTGCCGTATATGATGAAGGTGAGACGTCAAGTTGCCGGAATTTTAC	229		
Db 240	TATAGCATCAGACGAGATCCGGAATGTATGAAGGTGAGACGTACGTAGCCGGAATTTGAAC	299		
Qy 230	TTTCATTTGATTCATGCAATGTGGCGGTACACGATCTGTGATACAGCTGGTATTTTTCG	289		
Db 300	TTTCGTTTGATTCGTAATGTAGCAGCTACACGTTCCGTTAAATCCACCTGGTATTTTTCG	359		
Qy 290	TAAACAACAACGTTGTCAATTCGTTTCATTCACATTTATTTGATACCAAGTGTGATCGTCAT	349		
Db 360	TCACATCAAGTGTGTCTCATCTCATTCATCCATCTTTGTTGTGCAAAAAGTTGATCCAGCAT	419		
Qy 350	ATCGAGTCAATGCTTTTACATGGAAGCTGATPAAACAGTATAGTGCACAGATTGAGTAT	409		
Db 420	ATGGAATACATGCTTTTACATGGAAGCTGATPAAACAGTATAGTGCACAGATTGAGTAT	479		
Qy 410	CTGAATATCAACATGCTTTTCAAACTCAAAATGTCCCGATGCCAGTATCGGTTATGA	467		

					/strain="Kumba, Cameroons"
				/db_xref="taxon:6282"	
				/clone="SMOv3CAM51604"	
				/clone_1lp="Onchocerca volvulus molting L3 larva cDNA (SL96MLM-Ovml3)"	
				/dev_stage="molting L3"	
				/lab_host="Xil-Blue MRF"	
				/note="Vector: lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (mL3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by Rnase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10 <sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michale Litote-Wanilewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigm@nybc.org)."	
BASE COUNT	192 a	127 c	151 g	221 t	
ORIGIN					
Query Match	26.4%	Score 307;	DB 12;	Length 691;	
Best Local Similarity	77.6%;	Pred. No. 1e-74;			
Matches 463;	Conservative 0;	Mismatches 120;	Indels 14;	Gaps 7;	
OY	50	ATTGCATCCGGGTACACAATGCGTCGAAGGAGACCACCAAAATTTGAAATGGACCAACTT	109		
Db	82	ATGCTATTCGGTGTAATAACGGGTGAGAAGGGAACCGAATAATGTGGTCCAATTT	141		
OY	110	CAATPACAATAATTTTATAACAGCTAATGCAATTGGAAGACATGTTTATGGAAGGTC	169		
Db	142	CAATTACGTGCAACTTTTATACGTGTAATTCATTGGAAGGACATTTATACGGAAGGCT	201		
OY	170	TTTATGATCAAGAAGGTTGCCGTAAATGATGAAGGTGACGTCAAGTTCGCCGGAATTTTCA	229		

Db 202 TATACGATCAGGACGAGGACCCGAAATGATGAAGTGGACGTCAGTAGCCGGAAATTGAAC 261  
 Oy 230 TTCCATTGATTCATGATGATGGCGCGTACAGATCTCTGATTCACAGCTGGTATTTTG 289  
 Db 262 TTCCGTTGATTCGTTAATGATGACACGTTCCGTTAATTCACAGCTGGTATTTTG 321  
 Oy 290 TAACAACAAGCTGTGTGATTCGTTTCATTCATTTATTTGTAACAAAGTTGATTCGCAT 349  
 Db 322 TCACATCAGTGTGTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 381  
 Oy 350 ATCGATCAATGCTTTTACATGAGGATGAAGGATGAAGGATGAGGATGAGGATGAGGAT 409  
 Db 382 ATCGATCAATGCTTTTACATGAGGATGAAGGATGAAGGATGAGGATGAGGATGAGGAT 441  
 Oy 410 CTGAATACACAGCTGTTTCAACTCAATGAT--CCCGATGCCAGTATCCGTTATGAA 468  
 Db 442 TCCAAATGACAGCTGATTTTGACACAAATGTACCCGATCCGTTATCCGATATATGAG 501  
 Oy 469 ATT-TTGGATGGTGACCAACCGGTCAACC--AGTTCAATTTGCTATCATTT--GGTCAAC 523  
 Db 502 ATTCCTGGATGGTGACCAACCGGTCAACCCTGTTTCATTTGGTCAACATTTGGTCAACC 561  
 Oy 524 CAGTTATCATCAATGAAGACA-TGGATTCCTGAACCGTTGATA---CTTTCGCGCGGTT 579  
 Db 562 AAGTACCATTAATGACATGATGATTTCTGAACAGGTTGATACATTTCTGTGCTGTT 621  
 Oy 580 GTCCATTCTGCTTGTGTGATGAT---GGTAAACGCTGATGCTGTAATTTCTAA 632  
 Db 622 GTACAAATTCATGCTTCCGTTGGATGATGGGTAAACGCTGTAATGAGGAGGATTAATTA 678

RESULT 25  
 A1322068 552 bp mRNA linear EST 22-DEC-1998  
 LOCUS A1322068  
 DEFINITION SMOV3MCA12A04SK Onchocerca volvulus molting L3 larva cDNA  
 (SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA12A04 5',  
 mRNA sequence.  
 A1322068  
 VERSION A1322068.1 GI:4056219  
 KEYWORDS EST.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Onchocerca.

REFERENCE 1 (bases 1 to 552)  
 AUTHORS Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
 TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: Bluescript SK.

FEATURES  
 SOURCE Location/Qualifiers  
 1..552

/organism="Onchocerca volvulus"  
 /strain="Kumba, Cameroon"  
 /db\_xref="taxon:6282"  
 /clone="SMOV3MCA12A04"  
 /clone\_lib="Onchocerca volvulus molting L3 larva cDNA  
 (SL96MLM-Ovml3)"  
 /dev\_stage="molting L3"  
 /lab\_host="X1-Blue MRF"  
 /note="Vector: Lambda Uni-ZAP XR; Site.1: Eco RI; Site.2:  
 Xho I; Filarial nematode parasite of humans. Third-stage  
 larvae, L3, were isolated from infected black flies in  
 Cameroon (forest strain). The L3 were cultured in 20% FCS  
 in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in  
 culture. L3 of O. volvulus molt to fourth-stage larvae by

day 5 in culture. mRNA was isolated from approximately  
 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3  
 in culture, and converted to double-stranded cDNA using  
 reverse transcriptase and oligo(dT) followed by RNase H  
 and DNA pol I. The library was constructed in the lambda  
 Uni-Zap XR vector and has 1 x 10<sup>6</sup> independent  
 recombinants and the average insert size is ~1200 bp. The  
 library was constructed by Sara Lustigman and Michelle  
 Lizotte-Waniewski in the laboratory of Dr. S. A. Williams.  
 The library is available from Dr. Sara Lustigman (email:  
 slustigman@bc.org)."

BASE COUNT 154 a 106 c 111 g 181 t  
 ORIGIN  
 Query Match 26.4%; Score 306.8; DB 9; Length 552;  
 Best Local Similarity 83.4%; Pred. No. 1,1e-74;  
 Matches 372; Conservative 0; Mismatches 72; Indels 2; Gaps 2;

Oy 50 ATTCGATTCGCGTTGACATGATGGTGCAGAGGTGACCCGAAATGATGAGCAACTT 109  
 Db 105 ATGCTATTCGCGTTGATGATGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 164  
 Oy 110 CAATACATCAATTTTATATACCTAATGATTTGGAAGAGATGTTATGGAAGATC 169  
 Db 165 CAATACATCAATTTTATATACCTAATGATTTGGAAGAGATGTTATGGAAGATC 224  
 Oy 170 TTTATGATCAAGAGGTTGCCGTAATGATGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 229  
 Db 225 TTTATGATCAAGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 284  
 Oy 230 TTCCATTGATTCATGATGATGGCGCGTACAGATCTCTGATTCACAGCTGGTATTTTG 289  
 Db 285 TTCCGTTGATTCGTTAATGATGACACGTTCCGTTAATTCACAGCTGGTATTTTG 344  
 Oy 290 TAACAACAAGCTGTGTGATTCGTTTCATTCATTTATTTGTAACAAAGTTGATTCGCAT 349  
 Db 345 TCACATCAGTGTGTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 404  
 Oy 350 ATCGATCAATGCTTTTACATGATGAGGATGAAGGATGAAGGATGAGGATGAGGATGAGGAT 409  
 Db 405 ATCGATCAATGCTTTTACATGATGAGGATGAAGGATGAAGGATGAGGATGAGGATGAGGAT 464  
 Oy 410 -CTGAATACACAGCTGTTTCAACTCAATGAT--CCCGATGCCAGTATCCGTTATGAA 467  
 Db 465 CCCGAAATGACAGCTGATTTTGACACAAATGTGACGATCCGTTATCCGATATATGAG 524  
 Oy 468 AATTTGATGGTGACCAACCGGTTC 493  
 Db 525 GATTCGTGATGGTGACCAATCCCGGC 550

RESULT 26  
 A1317885 615 bp mRNA linear EST 17-DEC-1998  
 LOCUS A1317885  
 DEFINITION SMOV3MCA06H12SK Onchocerca volvulus molting L3 larva cDNA  
 (SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCA06H12 5',  
 mRNA sequence.  
 A1317885  
 VERSION A1317885.1 GI:4033152  
 KEYWORDS EST.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Onchocerca.

REFERENCE 1 (bases 1 to 615)  
 AUTHORS Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S.  
 TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA

Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: pBluescript SK.  
 Location/Qualifiers

# FEATURES

1. 615  
 /organism="Onchocerca volvulus"  
 /strain="Kumba, Cameroon"  
 /db\_xref="taxon:6282"  
 /clone="SMOV3KCM06H12"  
 /clone\_1lb="Onchocerca volvulus molting L3 larva cDNA (SL96MIM-Ovml3)"  
 /dev\_stage="molting L3"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda Uni-ZAP XR. Site\_1: Eco RI; Site\_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-ZAP XR vector and has 1 x 10<sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigman@hydc.org)."

BASE COUNT 172 a 124 c 123 g 189 t 7 others  
 ORIGIN

Query Match 25.8%; Score 299.6; DB 9; Length 615;  
 Best Local Similarity 82.5%; Pred. No. 1.le-72;  
 Matches 367; Conservative 0; Mismatches 75; Indels 3; Gaps 2;

50 ATTCGATTCGGTTGACAAAGTGTGAGGTGAGCAAGAAATGATGCGACCACTT 109  
 114 ATGCTATTCGGTTGACAAAGTGTGAGGTGAGCAAGAAATGATGCGACCACTT 173  
 110 CAATTAACAAATTTTAAATACAGTGAATGATGCAAGCAATGTTATGTGAAGGTC 169  
 174 CAATTAACAAATTTTAAATACAGTGAATGATGCAAGCAATGTTATGTGAAGGTC 233  
 170 TTTATGATCAAGAGTTCCTGATGATGAGGTGAGCAAGTTCGCGGAATTCAC 229  
 234 TATACGATCAAGGAGTTCCTGATGATGAGGTGAGCAAGTTCGCGGAATTCAC 293  
 230 TTCCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 289  
 294 TTCCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 353  
 290 TAACAAACA-CGTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 348  
 354 TCACATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 413  
 349 TATGAGTACAAATGCTTTTACATGAGAGTGTGAATAACA--GTTAGTGCACAGT 406  
 414 TATGAGTACAAATGCTTTTACATGAGAGTGTGAATAACA--GTTAGTGCACAGT 473  
 407 TATGAGTACAAATGCTTTTACATGAGAGTGTGAATAACA--GTTAGTGCACAGT 466  
 474 TTTCGAAATGACACAGTTCGATGACACAAATGATGACAGTTCGATGACAGTATG 533  
 467 AAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 491  
 534 AGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 558

RESULT 27  
 BE758436

LOCUS BE758436 543 bp mRNA linear EST 16-FEB-2001  
 DEFINITION SWACAL10F01SK Brugia malayi young adult cDNA (SAM99MLW-BmyA)  
 ACCESSION BE758436  
 VERSION Brugia malayi cDNA clone SWACAL10F01 5', mRNA sequence.  
 KEYWORDS EST.  
 SOURCE Brugia malayi.  
 ORGANISM Brugia malayi.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.

REFERENCE 1 (bases 1 to 543)  
 AUTHORS Williams, S.A.  
 TITLE Genes expressed in young adult of Brugia malayi  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: pBluescript SK.  
 Location/Qualifiers

1. 543  
 /organism="Brugia malayi"  
 /db\_xref="taxon:6279"  
 /clone="SWACAL10F01"  
 /clone\_1lb="Brugia malayi young adult cDNA (SAM99MLW-BmyA)"  
 /dev\_stage="young adult"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda Uni-ZAP XR. Site\_1: Eco RI; Site\_2: Xho I; Lymphatic filarial nematode parasite of humans. mRNA was prepared from young adult worms isolated from the peritoneal cavity of birds and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library has 6.5 x 10<sup>4</sup> independent recombinants and the average insert size is approx. 800bp. The library was constructed by Michelle Lizotte-Waniewski. The library is available from Dr. S.A. Williams, email: genome@neal.smith.edu."

BASE COUNT 140 a 119 c 123 g 159 t 2 others  
 ORIGIN

Query Match 25.8%; Score 299; DB 12; Length 543;  
 Best Local Similarity 80.0%; Pred. No. 1.le-72;  
 Matches 364; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

4 ATGATGCTTATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 63  
 90 ATGCTGCTTATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 149  
 64 GACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 123  
 150 GATACGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 209  
 124 TTTAATACAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 183  
 210 TTCAATACAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 269  
 184 GGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243  
 270 GGTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 329  
 244 TGCAATGTTGCGCTACACGATCTGATGATGATGATGATGATGATGATGATGATG 303  
 330 TGCAATGTTGCGCTACACGATCTGATGATGATGATGATGATGATGATGATGATG 389  
 304 GTCAATGTTGCGCTACACGATCTGATGATGATGATGATGATGATGATGATGATG 363  
 390 GTAAATTCCTTCATCCATGCTGCTGATGATGATGATGATGATGATGATGATGATG 449





```

/lab.host="DH10B"
/notes=vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
The library was constructed by Brandt Chapellin and Dr.
James McCarter at Washington University, St. Louis. The
cDNA was made by using Dynabead oligo-dT priming (Dyna1). The
PCR based library using a modified protocol from the
SMART PCR cDNA Synthesis Kit from Clontech. directionally
cloned into the UpG sites of pAMP1. Nematodes were
provided by Dr. David Bird and Daniel Snyder of North
Carolina State University."
BASE COUNT      158 a      122 c      111 g      143 t
ORIGIN

```

Query Match	25.2%;	Score 292.6;	DB 13;	Length 534;
Best Local Similarity	76.0%;	Pred. No. 9.8e-71;		
Matches 361; Conservative	0;	Mismatches 114;	Indels 0;	Gaps 0;

QY	403	GAGGATATCGAAATACACACGCTTTTCAATCGAAATGTCCGATGCAATATGCCG	462
Db	1	GGGGGTCCGAAATATACACACAAATCCAAACAAATTTGACTATGCTCTATGTAAA	60
QY	463	TATGAATTTTGGATGTGGACCAACCGGTCAACGATTCATTTGCTATCATTTGGTCAG	522
Db	61	TATGATTCCTCGAAGGGGGCCGACTGGTCAACCAATTCATTTGGCATTTATGGCCAA	120
QY	523	CCAGTTTTCATTAATGACATGGCATTTCTGAAACCGTTGATCTTCTGCGCGGTTGTC	582
Db	121	CAAGTTTATCAAAATGACCTTGATTTACAGGACAGTAGACACATTTGCGCTGTGTGA	180
QY	583	CATTTCCTGCTTGTGATGATGTGTAAACGGTGATCTGTGGAATAATCTGATGTA	642
Db	181	CATTCTGCTCTGTTGATGATGTGTATGGTGTACTGTCCAAATATTAAACCAAGAGGT	240
QY	643	TGTGCTCTTGATTAATTTGCTAAATATTGGAATATCCACAGATTTAATGGCTGGC	702
Db	241	TGTGCTTTGGATTAATTTCTTATTATAACCACTTGGAAATATCCAACTGATTTAATGGCTGTA	300
QY	703	CAAGAGCTCAAGTATTACAAATATCGATCGATCACAGCTTTCTATCAATGCCAGATC	762
Db	301	CAAGAGCTCAAGTATTATAAATATCTGATCTTACACAATTTTCTATCAATGCCAAAT	360
QY	763	AGTATTACCAATTAAAGAACCAAAATAGCGAGTGTTCGACCAACAATGTTCAGAACCA	822
Db	361	AGTATTACCAATTAAAGAACCTCACAGTGAATGTGCACGTCCAAAATGTGCAAGACTTCT	420
QY	823	GGATTCGGAGCGTAAAGAACAGGTGGTGGCCGCAAAAACGTCGACGACTCGCG	877
Db	421	GCGTTTAATGCTGTCAAAGTTGTGCTCTGGAGGGGGCGCGGACGAGAGCACACAC	475

RESULT 30	LOCUS	DEFINITION	AA901444
AA901444	531 bp	mrna	linear
AA901444	SMOVB3MCAM03B05	Onchocerca volutus molting L3 larva cDNA	EST 26-MAR-1999
AA901444	(SI:96KLW-Ovml.3)	Onchocerca volutus cDNA clone onch17 5' similar to	
AA901444	TR:Q93115 Q93115	CUT-1-LIKE CUTICLIN PROTEIN PRECURSOR. ; , mRNA	
AA901444	sequence.		
AA901444	AA901444		
AA901444.1	GI:3037198		
EST.			
Onchocerca volutus.			
Onchocerca volutus			
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;			
Onchocercidae; Onchocerca.			

REFERENCE  
AUTHORS  
1 (bases 1 to 531)  
Williams, S.A., Lizotte-Waniewski, M., Ianey, S., Instigman, S.,  
Hillier, L., Allen, M., Bowles, L., Gelsel, S., Jost, S., Kueba, T.,  
Mather, J., Stepien, M., Theising, B., White, Y., Wyllie, T., Chappell, J.,  
Person, B., Gibbons, M., Harvey, N., Pape, D., Chamberlain, A.,  
Morales, R., Schnur, R., Riter, E., Kohn, S., Underwood, K. and Marra,  
M.  
Molecular Parasitology Ovm13  
Unpublished (1998)

COMMENT

Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826

**FEATURES**

Email: [genomesmith.edu](mailto:genomesmith.edu)

The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S.A. Williams. The library is available from Dr. Sara Lustigman email [slustigman@bc.org](mailto:slustigman@bc.org) When requesting this clone from Dr. Lustigman please reference the Williams lab clone id - SMOY3MCAW03B05

Seq primer: 40m13 fwd. Et from Amersham

High quality sequence stop: 484.

Location/Qualifiers

```

/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:5282"
/clone="onch17"
/clone_1fb="Onchocerca volvulus molting L3 larva cDNA
(S196MM-Ovm13)"
/dev_stage="molting L3"
/lab_host="X11-Blue MRF"
/notes="Vector: Lambda Unit-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroon (forest strain). The L3 were cultured in 20% FCS
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (m13), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNA pol I. The library was constructed in the lambda
Uni-Zap XR vector and has 1 x 10E6 independent
recombinants and the average insert size is ~1200 bp. The
library was constructed by Sara Lustigman and Michelle
Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams
The library is available from Dr. Sara Lustigman (email:
slustig@nyc.corg)."

```

BASE COUNT 163 a 100 c 100 g 166 t 2 others

ORIGIN

Query Match 25.1% Score 291.2; DB 9; Length 531;  
 Best Local Similarity 83.5%; Pred. No. 2.4e-70;  
 Matches 329; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY	50	ATTGCATCCCGTGGACATGGTGTGCGAAGGAGACCCAGAAATTTAAATGTGACCAACTT	109
Db	137	ATGCTATTCCGGTGGATTAAGCGGTAGAAGGTGAACCAAGAAATTTGATGTGGTCCAAACT	196
QY	110	CAATTAACCAATCAATTTTAAATACGCTAATGCATTTGGAAGGACATGTTTATGTGAAAGATC	169
Db	197	CAATTAATCGTAACTTTTAATACGTGTAATCCATTTTGAAGGACATGTATATACGTGAAGGCT	256
QY	170	TTTATGATCAAGAAGGTGGCCGTATATGATGATGAAGTGTGACGTCNAAGTTGGCGGAATTTAC	229
Db	257	TATAGCATCAGGACGAGGATGCGNNATGATGAAGGTGAAGCTCAGGTAGCGGGAATTTGAC	316
QY	230	TTTCATTTGATTCATGCAATGTGGCGGTACACGATCTCTGAAATCCAGCTGATATTTTGG	289
Db	317	TTTCGTTGATTCGCTGATATGACGACGTACCGTTCGTTAAATCCACGTGGTATTTTGG	376
QY	290	TAAACACAACTGTGTGTCATTTGCTTCAATCCATTAATTTGTACCAAAGTTGATCGTCAT	349
Db	377	TCACATGAGTGTGTGTCATCTCATTTCCATCCATCTGTTGTGCAAAAAGTTGATCGAGCAT	436
QY	350	ATTCGATTCATAGCTTTTACATGGAAGCTGATAAACAAGTATGTGCACAGATTTGAGGTAT	409
Db	437	ATTCGAATTCATAGCTTTTACATGGAAGCTGATAAACAAGTATGTGCGCTCAACTTGAAGTTT	496



QY 410 CTGAATACACACTGCTTTTCAACTCAATGT 443  
 Db 497 CCGAATGACAACTGATTTGCAACAATGT 530

RESULT 31  
 B1783374  
 LOCUS 592 bp mRNA linear EST 26-SEP-2001  
 DEFINITION kh1b11.y1 Ascaris suum female gonad MZ PAMP1 v2 Chiapelli McCarter  
 Ascaris suum cDNA 5' similar to TR:093115 093115 CUF-1-LIKE  
 CUTICLIN PROTEIN PRECURSOR. ; mRNA sequence.

ACCESSION B1783374  
 VERSION B1783374.1 GI:15786266  
 KEYWORDS EST.  
 SOURCE Pig roundworm.  
 ORGANISM Ascaris suum  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea  
 ; Ascarididae; Ascaris.  
 1 (bases 1 to 592)  
 McCarter,J., Clifton,S., Chiapelli,B., Page,D., Martin,J., Wylie,T.,  
 Dante,M., Maira,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,  
 Gibbons,M., Rittler,E., Bennett,J., Franklin,C., Tsagarisvill,R.,  
 Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe  
 , M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,  
 Shin,T., Jackson,T., Cardenas,M., McCann,R., Waterston,R. and  
 Wilson,R.  
 The Washington Univ. Nematode EST Project, 1999  
 Unpublished (1999)  
 CONTACT McCarter JP  
 The Washington Univ. Nematode EST Project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 The library was constructed by Brandt Chiapelli and Dr. James  
 McCarter at Washington University, St. Louis. The cDNA was made by  
 using Dynabead oligo-dt priming (Dyna1). PCR based library using a  
 modified protocol from the SMART PCR cDNA Synthesis Kit from  
 Clontech. Directionally cloned into the UDG sites of PAMP1.  
 Dissected nematode tissues were provided by Dr. Alan Scott  
 (ascot@hsp.edu) of the School of Public Hygiene and Public Health  
 at John Hopkins University in Baltimore, MD.  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 389.  
 Location/Qualifiers  
 1..592  
 /organism="Ascaris suum"  
 /db\_xref="taxon:6253"  
 /clone\_lib="Ascaris suum female gonad MZ PAMP1 v2  
 Chiapelli McCarter"  
 /sex="Female"  
 /tissue\_type="Dissected female gonad (MZ-maturation zone,  
 most proximal region of gonad)"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B"  
 /note="Vector: PAMP1 (Gibco); Site\_1: NotI; Site\_2: SalI;  
 The library was constructed by Brandt Chiapelli and Dr.  
 James McCarter at Washington University, St. Louis. The  
 cDNA was made by using Dynabead oligo-dt priming (Dyna1).  
 PCR based library using a modified protocol from the SMART  
 PCR cDNA Synthesis Kit from Clontech. Directionally cloned  
 into the UDG sites of PAMP1. Dissected nematode tissues  
 were provided by Dr. Alan Scott (ascot@hsp.edu) of the  
 School of Public Hygiene and Public Health at John Hopkins  
 University in Baltimore, MD."

BASE COUNT 165 a 136 c 135 g 156 t  
 ORIGIN

Query Match 24.5%; Score 285; DB 13; Length 592;  
 Best Local Similarity 75.9%; Pred. No. 1.3e-68;  
 Matches 365; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 41 CATTGCTTATTCGATTCGCGTTGACATGCTGCAAGTGAGCCAGAAATGAATGTG 100  
 Db 99 CATTTCGTGTCGATCCCGGTTGATACGGCGTTAGGAGAGACCGGAATTTGAATGTG 158

QY 101 GACCACTTCATACAAATCAATTTTAATACAGTAAATGACATTCGAGACATGTTATG 160  
 Db 159 GCCCACTCAATTAACAGTCAACTTTAACACTGGAATCAATTTGAGGGGACGTTTACG 218

QY 161 TGAAGCTTTATATGATCAAGAAAGTGGCGTAAATGATGAAGTGAAGTCAAGTTCGCG 220  
 Db 219 TAAAGGACTCTATGATCAAGAAAGATGCGGTACATGAAGTGAAGTGAAGTTCGCG 278

QY 221 GAATTCACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 280  
 Db 279 GATACAGTTACCGTTCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 338

QY 281 GATATTTTGTAAACACACTGTTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 340  
 Db 339 GCATTCCTGACAGACGATTCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 398

QY 341 ATCGTCATATGAGATCAATGCTTTACATGAGACGATTAATGATTAAGTTCAGACAGA 400  
 Db 399 ATGACGCTATTCGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 458

QY 401 TTGAGGTATCTGAATACACACTGCTTTTCAACTCAATTCGATTCGATTCGATTCGATTC 460  
 Db 459 TGAAGTGTGCGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 518

QY 461 GTTATGAATTTTGGATGTTGATGACACACGCGTCAACGATTAATTTGCTATCATTTGTC 520  
 Db 519 GATACGAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 577

QY 521 A 521  
 Db 578 A 578

RESULT 32  
 BF400320  
 LOCUS BF400320  
 DEFINITION BF400320 622 bp mRNA linear EST 28-NOV-2000  
 SMOV3KCM44H03K Onchocerca volvulus molting L3 larva cDNA  
 (SI96KLM-ovm13) Onchocerca volvulus cDNA clone SMOV3KCM44H03 5',  
 mRNA sequence.  
 BF400320  
 BF400320.1 GI:11386388  
 EST.  
 ACCESSION BF400320  
 VERSION BF400320.1 GI:11386388  
 KEYWORDS EST.  
 ORGANISM Onchocerca volvulus.  
 Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Onchocerca.  
 1 (bases 1 to 622)  
 Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.  
 Genes expressed in molting L3 larvae of Onchocerca volvulus  
 Unpublished (1997)  
 JOURNAL  
 COMMENT  
 Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: Bluescript SR.  
 Location/Qualifiers  
 1..622  
 /organism="Onchocerca volvulus"  
 /strain="Kumba, Cameroons"  
 /db\_xref="taxon:6282"  
 /clone="SMOV3KCM44H03"  
 /clone\_lib="Onchocerca volvulus molting L3 larva cDNA  
 (SI96KLM-ovm13)"  
 /dev\_stage="molting L3"  
 /lab\_host="X1-Blue MRF"

BASE COUNT	189 a	118 c	135 g	180 t
ORIGIN				

Qy 466 GAATTTTGGATGCGTGACCAACCGGTCAACCAGTTCATTCTGCATCATTTGGTCAACCA 523  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2 GAGATTTCTGGAAGGCGACCAACTGTGACACCACTTCGATATGCAATGATGCGGAGATCAC 61

Dy 526 GTTATCATTAATGGACATCGATTCTGAAACCGTTGATACCTTTCGCGCGGTGTCCAT 583

Db 62 GATATCACCAATGGACTTGATTCAGAACACACAGATAATTTCTGCAATTAGTGCAT 121

Dy 586 TCCTGCTTGTGCAGATGGTAACGGTGATACTGTGGAATTCTAAATCGCATGGATGT 645  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 122 TCTTGACTGTGATGACGAAAAGTCATACGCTAGACATTCGAATGAAGAGCGTTGT 181

Dy      GCCTTGAAATAATTCGTAAATAATTGGAGATCCACAGATTTAATGCGTGCCTCA  
705

Db      GCGTGGACAAGTAATTACTCAACAATTTGGAAATCTCATCAGATCTAATGGCTGCCAA  
2411

27 100 GAGGCTCAGCTTTATAAATACGCTGATCGATTCGAGATTATTATTCATGCCAGATTAGT 3011

Db 302 ATCAACATTTAAAGAGCCATTGGCGAATGTCTTCGACCACATGTCCAGAACCACAAGA 361

Db 362 TTGGTGCATCAATCTGCTGCTGCCCAAGCTCCAGAGAGCTGCCCTACTTTACACACGA 421

Db 422 TTGCTCAAGACGATCGGTTAATTATGATAATACGGTGGATGTTAGTCCGGTTTACG 481

Db 482 GC GATTGAATTAAGCGAACAGATCCGACCTTTCAGCTAAT 523

RESULT 33	
AA625020	
LOCUS	
AA625020	363 bp
	mRNA
	linear
	EST 12-NOV-1997

(SL96MLW-OvML3) Onchoc  
sequence.  
AA625020

SOURCE	ORGANISM
Onchocerca volvulus.	Onchocerca volvulus
	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 363)	Williams, S.A., Likoite-Waniewski, M., Laney, S. and Lustigman, S.	Genes expressed in molting L3 larvae of <i>Onchocerca volvulus</i>	Unpublished (1997)	Contact: Steven A. Williams

Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826

FEATURES	
source	Email: genome@smith.edu Seq primer: pbluescript SK. location/Qualifiers 1. .363

```

/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone="SmmJ3C0591"
/clone_lib="Onchocerca volvulus molting L3 larva cDNA"

```

/note=Vector: Lambda Uni-Zap XR; Site:1: Eco RI; Site:2: Xho I; Filariarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of *O. volutus* molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda uni-Zap XR vector and has 1 x 10<sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizzotte-Waniewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigman@cbc.org)."

BASE COUNT	104 a	71 c	77 g	110 t	1 others
ORIGIN					

Query Match	24.0%;	Score 278.8;	DB 9;	Length 363;
Best Local Similarity	85.4%;	Pred. No. 6e-67;		
Matches 310; Conservative	0;	Mismatches 53;	Indels 0;	Gaps 0;

[illegible]

Db 61 GCCTGTATGCCGATTGAGATTCTTGATGGTGCACCATCCGGACAACCTGTTCAATTTC 120

D6 121 TACCATTCGTCACCAAGTGCATCATTAATGGACATGTGATTCTGAACACAGTTGATACATT 180

Db 181 CTGTCGTGTACATTCATCCTTCGTGATGATGGTAACGGTGATTAAGGTGGAATTATT 240

241 AAATGCTGATGTTGTGCTTTGGACAATAATTGCTCAACAACITGGAAATATCCAACTGA 300

Db	301	TTTATGGCTGGACAGAAGCTCAGCTCTACANTATGGGATGATCACAACTGTTCTA	360
Oy	750	TCA	752

```

Db      361 CCA 363

RESULT 34
AM313040
LOCUS   558 bp      mRNA      linear      EST 24-JAN-2000
DEFINITION
SMOV3M3CAM38A02SK Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3M3CAM38A02 5',
mRNA sequence.
ACCESSION
AM313040
VERSION
AM313040.1 GI:6742215
KEYWORDS
EST.
SOURCE   Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocerca.
REFERENCE
1 (bases 1 to 558)
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
TITLE    Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL  Unpublished (1997)
COMMENT  Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
FEATURES
source
1..558
location/Qualifiers
/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone="SMOV3M3CAM38A02"
/clone_1lb="Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3)"
/dev_stage="molting L3"
/lab_host="XLI-Blue MRF"
/notes="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroon (forest strain). The L3 were cultured in 20% FCS
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNA pol I. The library was constructed in the lambda
Uni-zap XR vector and has 1 x 10E6 independent
recombinants and the average insert size is ~1200 bp. The
library was constructed by Sara Lustigman and Michelle
Lizotte-Waniewski in the laboratory of Dr. S. A. Williams.
The library is available from Dr. Sara Lustigman (email:
slustig@mcnyc.org)."
BASE COUNT  167 a 111 c 114 g 163 t 3 others
ORIGIN
Query Match      23.9% Score 277.8; DB 10; Length 558;
Best Local Similarity 71.9%; Pred. No. 1,4e-66;
Matches 388; Conservative 0; Mismatches 150; Indels 2; Gaps 2;

QY 7 ATTCGCTTAATGCTTCTGTAACACTATTCATTGATTCGATTCGGGTGA 65
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 17 ATCTCTCTCTGCTAATGTAATGTAATGTAATGTAATGTAATGTAATG 76
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 66 CAATGCTGTGGAAGTGAAGCAAAATGATGTGACCAACTCAATACATCAATT 125
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 77 TAACGGTGTGCAAGTGAACCTGAATGAATGATGACCACTGCTCAACATCAATT 136
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 126 TATATACGTAATGCAATGTGGAAGACATGTTTATGTGAAGGCTTTATGATCAACAGAG 185
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db      137 CAATACATAGAAATACATTTGAGGTGATGATACGTAAAGGCTTTACATGCGCATGA 196
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 186 TTGCCATATGATGAGAGGTGAGAGTCAAGTTGCCGGAATTTCACTTCATTCATCAG 245
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 197 ATGTCGTTCAAGATAGTGTGAGCGGAGTACGATGAATGAATGATGATGATGATGATG 256
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 246 CAATGTTGGCCGTACAGATCTGATGATCCAGCGTGTATTTTGTACCAACAACATGTTGT 305
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 257 TAATGTTGAAGCATACAGAGGTCCTTAATCTCGAGGTCCTTTGTATCAACCATTAATG 316
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 306 CATTTGCTTTCATTCATTCATTTATTTGTTACCAAGTTGATCGTCAATACGATACATGCTT 365
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 317 TATTCATTTTCATCCGAAATTCATTTCAAAAATGAGACCGCATTCATGATACAGTCTTT 376
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 366 TTACATGAGAGCTGATAAACAGTTGTCAGCATTTGAGCTATGCAATCAACATCAG 425
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 377 CTATATGAGAGCTGACAAACCGTTGATGAGACTTGAAGTATCCGAATGACACAGC 436
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 426 TTTTCAAACTCAAAATTTGTCCTGATGCGCAGTATGCCCTTATGAATTTTGGATGTGAGCC 485
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 437 ATTCAGACACAGAGTGTATCAATGCTGTGTGATCCAGAAATCTGGAANGNGAGC 496
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 486 AACCGGTCAACAGCTTCAATTTCTATCATTTGTCAGC-CAGTTTATCATTAATGACAT 544
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 497 AACTGTACACCTATTGATGATGCAATGATGAGATACACCGTATATCACAAAATGGGCT 556
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 35
AM312995
LOCUS   543 bp      mRNA      linear      EST 24-JAN-2000
DEFINITION
SMOV3M3CAM37E11SK Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3M3CAM37E11 5',
mRNA sequence.
ACCESSION
AM312995
VERSION
AM312995.1 GI:6742181
KEYWORDS
EST.
SOURCE   Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocerca.
REFERENCE
1 (bases 1 to 543)
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
TITLE    Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL  Unpublished (1997)
COMMENT  Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
FEATURES
source
1..543
location/Qualifiers
/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone="SMOV3M3CAM37E11"
/clone_1lb="Onchocerca volvulus molting L3 larva cDNA
(SL96MLM-Ovml3)"
/dev_stage="molting L3"
/lab_host="XLI-Blue MRF"
/notes="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroon (forest strain). The L3 were cultured in 20% FCS
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using

```

BASE COUNT	160 a	104 c	110 g	166 t	3 others
ORIGIN					

0y	ATTGCAATTCGGGTGACAAATGCTGTGCAAGGAGACCCAGAAATTCGAATGGACCAACT	109
122	ATGCAATTCGGGTGATTAACGGTGTAGAAGGTGAACCCAGAAATTCGAATGGTCCAAACT	181
0y	CAATACAAATCAATTTAATACAGTAAATGCAATTCGGAAGGACATGTTATATGGAAGATC	169
182	CAATTAACGTCAACTTAAATCTGTAATCCATTTGGAAGGACATGTTATACGGAAGGCT	241
0y	TTTATGATCAAGAGCGTTCCGTAATGATGAAGGTGGACGTCAAATGTCCGGAATTTCC	229
242	TATACGATCCAGGACGATATCCGAATATGATGAAGGTGGACGTCGATAGCGGGAATTTGAC	301

[illegible]

RESULT 36	LOCUS	DEFINITION
AM288114	577 bp	linear
AN288114	577 bp	linear
SMOV3MCAM33H10SK	Onchocerca volvulus molting L3 larva cDNA	
(SI596ILM-Ovnm2.3)	Onchocerca volvulus cDNA clone SMOV3MCAM33H10 5', mRNA sequence.	

ACCESSION	AM288314	
VERSION	AM288314.1	GI:6695049
KEYWORDS	EST	
SOURCE	<i>Onchocerca volvulus</i>	
ORGANISM	<i>Onchocerca volvulus</i>	
	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Onchocerca.	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 577)	Williams, S.A., Lizotte-Wanlewska, M., laney, S., and Lustigman, S.	Genes expressed in molting L3 larvae of <i>Onchocerca volvulus</i>	Unpublished (1997)	Contact: Steven A. Williams

**COMMENT**  
Contact: Steven A. Williams  
Molecular Parasitology  
Smith College department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA

```

Email: genomesmith.edu
Seq primer: plbluescript SK.
Location/Qualifiers
1. 577
/organism="Onchocerca volvulus"

```

```

/strain="Kumba, Camereroons"
/db_xref="taxon:5282"
/clone="SMO03WCAM3H10"
/clone_1id="Onchocerca volvulus molting L3 larva cdna
(SL96MIM-Ovml3)"
/dev_stage="molting L3"
/lab_host="XLI-Blue MRF"
/note="Vector: lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroun (forest strain). The L3 were cultured in 20% FCS
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
culture. L3 of O. volvulus molt to fourth-stage larvae by
day 5 in culture. mRNA was isolated from approximately
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3
in culture, and converted to double-stranded cDNA using
reverse transcriptase and oligo(dt) followed by RNase H
and DNA pol I. The library was constructed in the lambda
Uni-Zap XR vector and has 1 x 106 independent
recombinants and the average insert size is ~1200 bp. The
library was constructed by Sara Lustigman and Michelle
Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
The library is available from Dr. Sara Lustigman (email:
slustigmen@bc.org)."

```

Query Match	23.7%;	Score 274.6;	DB 10;	Length 577;
Best Local Similarity	-82.4%;	Pred. No. 1.1e-65;		
Matches 327; Conservative	0;	Mismatches 69;	Indels 1;	Gaps 1;

QY 50 ATTCGATTCGGGTGACATGTTGCAGGAAGGACCAGAAATTGAATGTGAGCCACTT 109  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 162 ATGCAATTCGGGTTCATTAACGGTGTAGAAGGTGAACCGAATAATTAATGTGTCCAAAT 221  
  
QY 110 CAATAACAATCAATTTTAATACAGCTAATGCAATTCGAAGAACATGTTATGTGAAGAAGTCC 169  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 222 CAATTACTGTCAACTTTTAATCTGTATTCATTTTGAAAGACATGTATACGTGAANAAGCT 281  
  
QY 170 TTTATGATCAGAAGGTTCCCGTATATGATGAAGGTGGACGTCAAGTTGCCGAATTTTAC 229  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 282 TATACGATCACAGCAGGAATGCCGAATGTAGAAGGTGACGTCAAGTTACCGGAATTTGAAC 341  
  
QY 230 TTCAATTTGATTCATGACATGTTGGCGCTAACACGATCTCGAATCCAGGTGATTTTTG 289  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 342 TTCGGTTGATTCCTGTAATGAGACCATPACCGTTCCTTAATCCAGGTGATTTTTG 401  
  
QY 290 TAACAACAACAGTGTGTCAATTCGTTTCATCCATTAATTTGTTACCAAAGTGTGATCGTCAT 349  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 402 TCACATCAGTTGTSTCATCTCATTCATTCATTCATTTTGGGACAANAATTGATGCAGAT 461  
  
QY 350 ATCGAGTCAATGCTTTTACATGCAAGCTGTATAAACAAGTTAGTCACAGATTTGAGGTAT 409  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 462 ATCGAAT-CAATGCTTTTACATGCAAGCTGTATAAACAAGTTAGCGCTCAACTTTGAAAGTTT 520  
  
QY 410 CTGAATCAACAAGCTTTTCAANCTCAAAATGTGCC 446  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 521 CCGAATTCGCACTGGCAATTTGCAADACCAAAATGTACC 557

RESULT 37	
AA841200	
LOCUS	354 bp mRNA linear EST 02-MAR-1998
DEFINITION	M3B56AA461073 Brugia malayi day 6 post-infection third stage larvae
	5M956M1W-Bml3d6 Brugia malayi CDNA clone 3D6AA4G10 5', mRNA
	sequence.
ACCESSION	AA841200
VERSION	AA841200.1 GI:2922536
KEYWORDS	EST.
SOURCE	Brugia malayi.
ORGANISM	Brugia malayi
	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
	Onchocercidae; Brugia.

REFERENCE 1 (bases 1 to 354)  
AUTHORS Blaxter,M.L., Waterfall,M., Daub,J., Lizotte-Waniewski,M., Baron,L.  
and Jones,S.J.  
TITLE Genes expressed in day six post-infection, third stage larvae of  
Bugia malayi  
JOURNAL Unpublished (1997)  
COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The ABI trace of this sequence can be viewed at  
http://www.sanger.ac.uk/Bugia/3D6/MB3D6AA4G10T3.html  
Seq primer: T3.

FEATURES  
source  
1..354  
Location/Qualifiers  
/organism="Bugia malayi"  
/strain="T3S Labs"  
/db\_xref="taxon:6279"  
/clone="3D6AA4G10"  
/clone.lib="Bugia malayi day 6 post-infection third stage  
larvae SAW96MLW-Bml3d6"  
/sex="mixed"  
/dev\_stage="third stage larvae, six days after infection"  
/lab\_host="E. coli XL1-Blue"  
/note="Vector: lambdaZapRI (Unizap XR); Site:1: Eco R I  
(5' end); Site:2: Xho I (3' end); Bugia malayi is a  
lymphatic filarial nematode parasite of humans. mRNA was  
prepared from third stage larvae of Bugia malayi isolated  
from the peritoneal cavity of jirds six days after  
infection. The mRNA was converted to double stranded cDNA  
using reverse transcriptase and oligo(dT) followed by  
Rnase H and DNAPol I. The library had 2 x 10E5 independent  
recombinants and average insert size was 900 base pairs.  
The library was constructed by Michelle Lizotte-Waniewski.  
The library is available from The Filarial Genome Project  
Resource Center: contact Dr. S.A. Williams, Clark Science  
Center, Smith College, Northampton, MA 01063 USA phone +1  
413 585 3826 fax +1 413 585 3786 email genome@smith.edu."

BASE COUNT 105 a 81 c 84 g 84 t

ORIGIN

Query Match 23.2%; Score 269.2; DB 9; Length 354;  
Best Local Similarity 85.0%; Pred. No. 2.9e-64;  
Matches 301; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 638 ATGGATGCTCTTGAATTAATTTGCTAATATTTGGCAATATCCACAGATTAAATG 697  
|||||  
DB 1 ATGGTGTGCTCGACAAATATTTGCTAATATTTGGCAATATCCACAGATTAAATG 60  
QY 698 CTGGCCAGAAGCTCAGTATACAAATATGCGATGATCAGCTTTCTATCATATGCC 757  
|||||  
DB 61 CTGTGTAAGAGGCGCATGTGTACAAATATGCGATGATCAGCTTTCTATCAATATGCC 120  
QY 758 AGATCAGTATTAACATAAGAACCAATAGCAATGTGTCACACATATGTCAGAAC 817  
|||||  
DB 121 AGATCAGTATTAACATAAGAACCAATAGCAATGTGTCACACATATGTCAGAAC 180  
QY 818 CACAGAGATTGAGCTGTTAAACAGGTGTCGCCAGCAAAACCTGCTGCAGTCCG 877  
|||||  
DB 181 CCGAAGATTGGAGCTGTTAAACAGGTGTCGCCAGTAACTGCAGAGCGTGC 240  
QY 878 AACTTGTTTACTCAAGAAAAGATTCGCAACCGGAGATATCATGTATGTAAGACTG 937  
|||||  
DB 241 AACTCGCTGCTCGAAGAAAAGTTCGCGAACCAGAACGTCGTCGTCAGAGACTG 300  
QY 938 ATATCAACACCTTGAAATTAAGGATGATATCAAGCTTTGCGAGTTGATTTAC 991  
|||||  
DB 301 ATATCAGCCCTTGAAATTAAGGATGATATCAAGCTTTGCGAGTTGATTTAC 354

RESULT 38  
BF918256  
LOCUS  
DEFINITION SMOV3MCAM53B02SK Onchocerca volvulus molting L3 larva cDNA  
(SI96MLM-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCAM53B02 5',  
mRNA sequence.  
ACCESSION BF918256.1 GI:12213935  
VERSION BF918256  
KEYWORDS EST.  
SOURCE Onchocerca volvulus.  
ORGANISM Onchocerca volvulus.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Onchocerca.  
REFERENCE 1 (bases 1 to 705)  
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.  
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus  
JOURNAL Unpublished (1997)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: pluescript SK.

FEATURES  
source  
1..705  
Location/Qualifiers  
/organism="Onchocerca volvulus"  
/strain="Kumba, Cameroons"  
/db\_xref="taxon:6282"  
/clone="SMOV3MCAM53B02"  
/clone.lib="Onchocerca volvulus molting L3 larva cDNA  
(SI96MLM-Ovml3)"  
/dev\_stage="molting L3"  
/lab\_host="XL1-Blue MRF"  
/note="Vector: Lambda Uni-Zap XR; Site:1: Eco RI; Site:2:  
Xho I; Filarial nematode parasite of humans. Third-stage  
larvae, L3, were isolated from infected black flies in  
Cameroon (forest strain). The L3 were cultured in 20% FCS  
in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in  
culture. L3 of O. volvulus molt to fourth-stage larvae by  
day 5 in culture. mRNA was isolated from approximately  
6000 molting larvae (ml3), 2000 larvae from day 1, 2 or 3  
in culture, and converted to double-stranded cDNA using  
reverse transcriptase and oligo(dT) followed by Rnase H  
and DNA pol I. The library was constructed in the Lambda  
Uni-Zap XR vector and has 1 x 10E6 independent  
recombinants and the average insert size is ~1200 bp. The  
library was constructed by Sara Lustigman and Michelle  
Lizotte-Waniewski in the laboratory of Dr. S. A. Williams.  
The library is available from Dr. Sara Lustigman (email:  
slustigmen@bc.org)."

BASE COUNT 195 a 139 c 152 g 217 t

ORIGIN

Query Match 23.2%; Score 269.2; DB 12; Length 705;  
Best Local Similarity 82.2%; Pred. No. 3.7e-64;  
Matches 332; Conservative 0; Mismatches 70; Indels 2; Gaps 2;

QY 50 ATTCCATTCGCGTGTGACATGCTGCAAGGTGACGCCAGAAATTAATGATGGACCACTT 109  
|||||  
DB 117 ATGCTATTCGCGTGTGATTAACGCTGTGAGAGGAGAACCGAAATTAATGCTGCCAAT 176  
QY 110 CAAATACAAATCAATTTAATACACGTAATGCAATTCGAAGGACATGTTATGTGAAGCTC 169  
|||||  
DB 177 CAATTAAGTGTCAACTTAATCTGTATTCATGTAAGGACATGATATAGCGAAAGCT 236  
QY 170 TTATGATCAGAGAGTTCGCTATATGTAAGGTGAGCGTCAGTCCGGAATTTAC 229  
|||||  
DB 237 TATACGATCAGCAGAGATGCGGAATGATGAAGGTGAGCGTACAGTACCGAAATTTGAC 296

QY	230	TTCCATTTGATTCATGCAATGTTGGCGCTACACCAATCTGTGAATCCAGTGGTATTTTGG	289
Db	237	TTTCCTTTGATTCCTGTAATGAGACGACGCTTCGTTAAATCCAGTGGTATTTTGG	356
QY	290	TAAACAACAAGTGTGCATATTTGCTTCATCCATTTATTTGTATGCCAAGTGTATCGTCAT	349
Db	357	TCACATCAGTGTGTGTCTACATCCATTCATCCATTT -CTTGTGCAAAAAGTGTATCGAGTA	415
QY	350	ATCGAGTCAATGCTTTTACATGGAAGCTGATTAACAAGTATGTCACAGATT -GAGGTA	408
Db	416	ATCGAATTCATGCTTTTACATGNAAGCTGATAAACAAGTATGCGCTCAACTTGGAAATT	475
QY	409	TCTGAATCCACAGCTGCTTTTCAACCTCAAAATGTGCCGATGCC	452
Db	476	NCCGAATGACACAGCTGCTTTGCAACAACAATTTGATCCGATGCC	519

```

RESULT 39
AA625024
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA625024 466 bp mRNA linear EST 12-NOV-1997
SMOV3KCAV58SK Onchocerca volvulus molting L3 larva cDNA
(ST936MLM-Ovmlm3) Onchocerca volvulus CDNA clone SMML3CO758 5', mRNA
sequence.
AA625024
AA625024.1 GI:2537426
EST.
Onchocerca volvulus.
Onchocerca volvulus
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocera.
1 (bases 1 to 466)
Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
Genes expressed in molting L3 larvae of Onchocerca volvulus
Unpublished (1997)
Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
Location/Qualifiers
1..466

```

BASE COUNT	ORIGIN
126 a	91 c 93 g 144 t 12 others

Query Match	23.1%:	Score 267.8;	DB 9;	Length 466;
Best Local Similarity	80.7%;	Pred. No. 7.8e-64;		
Matches 305;	Conservative 0;	Mismatches 73;	Indels 0;	Gaps 0;
QY	ATTGCATTCGCGTGGACATGCTGTGCAAGGTGAGCCAGAAATTGAATGGAGCCAACTT	109		
Db	89 ATGCGATTCCGGTTGATTAACGGTTAGAAAGGTGAACCGAAATTGAATGTGGTCCAACT	148		
QY	110 CAATPACATCAATTTTAATACAGTAATGCATTTGCAAGACATGTTTATGTGAAAGCTC	169		
Db	149 CAATTACTGTCGAACCTTTTAATCTGTAATCCATTGGAAGGACATTTTACGTGAAAGCT	208		
QY	170 TTTATGATCAAGAAAGGTGGCCGTATGTGATGAAAGGTGAGCGTCAAGTTTCCGGAATTTGAC	229		
Db	209 TATACGATCAGGACGAGATCCGGAATGATGAAGGTGAGCGTGAAGTTCGGAATTTGACC	268		
QY	230 TTCCATTTGATTCATGATGCAATGTTGGCGGTACACGATCTCTGAATTCACGATGGTATTTTGG	289		
Db	269 TTCCGATGATTCCTGTAATGTAGACGCTTACACGTTCCGTANNKTCACGATGGTATTTTGG	328		
QY	290 TAAACAACAAGTGTGTCATTTGTTTCATTCATATTTTGTATACCAAATTTGATCGTGCAT	349		
Db	329 TCACATTCAGTTGTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	388		
QY	350 ATCGCATCAATTCGCTTTTACATGGAAGCTGATTAACCAAGTATAGTGCACAGATTTAGAGTAT	409		
Db	389 ATCGCATCAATTCGCTTTTACATGGAAGCTGATTAACCAAGTATAGTGCACAGATTTAGAGTAT	448		
QY	410 CTGAATTCACACATGCTTT 427			
Db	449 CCGAATATGACCAATGTCAT 466			

RESULT	40
LOCUS	AA618952
DEFINITION	AA618952 459 bp mRNA linear EST 12-NOV-1997
ACCESSION	AA618952
VERSION	SMOVCACA0205K
KEYWORDS	Onchocerca volvulus molting L3 larva cDNA (S156MLM-Ovml3) Onchocerca volvulus cDNA clone SMLM3CO2020 5', mRNA sequence.
SOURCE	AA618952.1 GI:2522828
ORGANISM	EST. Onchocerca volvulus. Onchocerca volvulus Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Onchocerca. 1 (bases 1 to 459) Williams, S.A., Lizotte-Waniewski, M., Laney, S. and Lustigman, S. Genes expressed in molting L3 larvae of Onchocerca volvulus unpublished (1997) Contact: Steven A. Williams Molecular Parasitology Smith College Department of Biological Sciences Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA Tel.: 4135853826 Fax: 4135853786 Email: genome@smith.edu Seq primer: Bluescript SK. Location/Qualifiers 1..459
FEATURES	
SOURCE	

```

/note=Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. Third-stage
larvae, L3, were isolated from infected black flies in
Cameroon (forest strain). The L3 were cultured in 20% FCS

```

in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of *O. volvulus* molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (ML3), 2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10<sup>6</sup> independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigmen@ybc.org)."

BASE COUNT 131 a 85 c 91 g 150 t 2 others  
ORIGIN

Query Match 22.7%; Score 263; DB 9; Length 459;

Best Local Similarity 83.9%; Pred. No. 1.7e-62;

Matches 296; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
QY 50 ATTGATTCGGGTGACAAATGCTGCGAGGTGACCCAGAAATTTGAATGTGACCAACTT 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 ATGCTATTCGGTGTGATTAAGGTGATGAGGTGAAACCAAAATTTGAATGTGTCACACAT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 110 CAATTAACATCAATTTTAATACAGTAATGATTCGAAAGACATGTTATGTGAAGTTC 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 CAATTAAGTCAACTTTAATACGTAATTCATTTGAGGACNTGTATACGTGAAGGCT 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 TTTATGATCAAGAGGTGGCGTAATGATGATGAGGTGACGCTCAAGTTCGCGAATTTT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 TATACGATCAGCAGGATGCCGAATGATGAGGTGACGCTCAAGTTCGCGAATTTT 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 TTCCATTTGATTCATGCAATGCTGCGGTACAGATCTGAAATCCACGCTGATTTT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 TTCCGTTTGTCTCTGAATGTANCAAGTTCAGCTTAAATCCACGCTGATTTT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 TAACAACAACCTGTTGCAATTCCTTCAATTTTGTACCAAAAGTGAATGCGTCAT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 TCACATCAGTGTGTCATCTCATTCATTCATTTGTTGACCAAAAGTGTGATCGAGCAT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 ATCGAGTACAACTTTTACATGGAAGCTGATAAACAGTTAGTGACAGATT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 ATCGAATACAACTTTTACATGGAAGCTGATAAACAGTTAGTGACAGATT 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 28, 2003, 16:58:33  
Job time : 1655 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 08:49:12 ; Search time 41 seconds  
(without alignments)  
1257.756 Million cell updates/sec

Title: US-10-054-562a-4

Perfect score: 2026

Sequence: 1 MMRLIAFCTTLIALSYSP.....ALIAAVITTSFRRPNQKA 387

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A.Geneseq\_101002.\*  
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2026	100.0	387	22	AA884258	Amino acid sequenc
2	2026	100.0	387	23	AB876278	Dirofilaria immiti
3	1086.5	53.6	245	22	AA384260	Amino acid sequenc
4	1086.5	53.6	245	23	AB876280	Brugia malayi cuti
5	776.5	38.3	271	22	AA884259	Amino acid sequenc
6	776.5	38.3	271	23	AB876279	Dirofilaria immiti
7	199	9.8	462	22	AB859975	Drosophila melanog
8	166	8.2	604	22	AB870061	Drosophila melanog
9	163	8.0	1638	22	AB867889	Drosophila melanog
10	162.5	8.0	611	22	AB867239	Drosophila melanog

11	162.5	8.0	611	22	AB870054	Drosophila melanog
12	150	7.4	758	22	AB858337	Drosophila melanog
13	142	7.0	833	22	AB858185	Drosophila melanog
14	140	6.9	632	22	AB870355	Drosophila melanog
15	135.5	6.7	744	22	AB858426	Drosophila melanog
16	130.5	6.4	638	22	AB864380	Drosophila melanog
17	124.5	6.1	699	22	AB864367	Drosophila melanog
18	119.5	5.9	418	22	AB868164	Drosophila melanog
19	111	5.5	2284	22	AB871444	Drosophila melanog
20	100.5	5.0	768	22	AB868660	Drosophila melanog
21	100	4.9	935	13	AA820112	AE-III (peptidylhy
22	97	4.8	227	22	AB871814	Drosophila melanog
23	94.5	4.7	633	10	AA894856	Expression plasmid
24	94.5	4.7	875	10	AA894854	C-terminal prepro-
25	94.5	4.7	875	16	AA873053	Peptidyl C-termina
26	93	4.6	474	22	AB863130	Drosophila melanog
27	92	4.5	896	23	AB808760	Synechococcus cyan
28	90	4.4	2515	22	AB864427	Drosophila melanog
29	87.5	4.3	474	21	AA870900	Protein encoded by
30	87.5	4.3	474	21	AA870923	Human soluble CD39
31	87	4.3	473	21	AA870901	Protein encoded by
32	87	4.3	473	21	AA870924	Human soluble CD39
33	87	4.3	3038	16	AA874171	Aspergillus terreu
34	86.5	4.3	636	22	AA894352	Human protein sequ
35	86.5	4.3	747	22	AA840070	Human polypeptide
36	86.5	4.3	747	22	AA893124	Human protein sequ
37	85.5	4.2	438	23	AB838284	Staphylococcus epi
38	85.5	4.2	560	18	AA813009	Segment of desmoso
39	85.5	4.2	940	22	AA839436	Human polypeptide
40	85.5	4.2	1117	23	AA878053	Human desmoglein 2
41	85.5	4.2	1118	23	AA820266	Human lung specif
42	85.5	4.2	1120	22	AB810284	Novel human diagno
43	85.5	4.2	1121	22	AA841282	Human polypeptide
44	85	4.2	2069	22	AA830566	C glutamicum prote
45	84	4.1	474	22	AA894481	Human protein sequ

#### ALIGNMENTS

RESULT 1  
AA884258  
ID AAB84258 standard; Protein: 387 AA.  
XX  
AC AAB84258;  
XX  
DT 22-AUG-2001 (first entry)  
XX  
DE Amino acid sequence of a cuticlin polypeptide.  
XX  
KW Cuticlin; gene therapy; vaccine; helminth parasite.  
XX  
OS Dirofilaria immitis.  
XX  
PN US6248329-B1.  
XX  
PD 19-JUN-2001.  
XX  
PF 01-JUN-1999; 99US-033427.  
XX  
PR 01-JUN-1998; 98US-0087435.  
XX  
PA (CHAN/) CHANDRASHEKAR R.  
XX (MORA/) MORALES T H.  
XX Chandrashekar R, Morales TH;  
XX WPI, 2001-396953/42.  
XX DR N-PSDB; AAH24826, AAH24828.  
XX  
PT Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin  
PT protein is useful as a vaccine to prevent parasitic helminth infection

XX Claim 3: Column 39-40; 29pp; English.  
PS  
XX  
CC The specification describes a Dirofilaria immitis nucleic acid molecule,  
CC that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid  
CC molecule is useful as a probe to identify nucleic acid molecules, as a  
CC primer to produce nucleic acids, as a therapeutic reagent, e.g., gene  
CC therapy to inhibit cuticlin activity or production, or in a vaccine to  
CC prevent infection with helminth parasites. The cuticlin protein,  
CC antibodies raised against it, and inhibitory compounds of cuticlin may  
CC all be used in compositions to protect animals, especially mammals such  
CC as cats, dogs, and humans. The antibodies may be used to passively  
CC immunize an animal, or as reagents in assay to detect infection of  
CC helminths, or as tools to screen expression libraries to recover desired  
CC proteins. They may also be used to target cytotoxic agents to the  
CC parasite and kill it directly. The present sequence represents a cuticlin  
CC polypeptide.  
XX  
SO Sequence 387 AA:  
Query Match 100.0%; Score 2026; DB 22; Length 387;  
Best Local Similarity 100.0%; Pred. No. 7.1e-210;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMRLIAFCTTLLALSYSIPVNGVGEPEIECGPTSTINFNTRNAFEGHYVVGGLYDQ 60  
DB 1 MMRLIAFCTTLLALSYSIPVNGVGEPEIECGPTSTINFNTRNAFEGHYVVGGLYDQ 60  
QY 61 ECGRNDEGGRQVAGISLPDSCNVARTRSINPRGIFVTTTVVISHPLFVTVDRAYRQ 120  
DB 61 ECGRNDEGGRQVAGISLPDSCNVARTRSINPRGIFVTTTVVISHPLFVTVDRAYRQ 120  
QY 121 CFYMEADKTVSAQIEVSETTAFQTQIYMPVCRYEILDGGPTGQVQFAIIGQPYVHKM 180  
DB 121 CFYMEADKTVSAQIEVSETTAFQTQIYMPVCRYEILDGGPTGQVQFAIIGQPYVHKM 180  
QY 181 TCDSEVDFFCVAVHSCFVDDNGDVTVEILNADGALDKYLLNNLEYPTDLMAAGEAHY 240  
DB 181 TCDSEVDFFCVAVHSCFVDDNGDVTVEILNADGALDKYLLNNLEYPTDLMAAGEAHY 240  
QY 241 KYADRSQLFYQCOISITTIKEPNSCEVRPQCSPOGFGAVKTGGAAPKAPAAAQRLKKR 300  
DB 241 KYADRSQLFYQCOISITTIKEPNSCEVRPQCSPOGFGAVKTGGAAPKAPAAAQRLKKR 300  
QY 301 SAEPENIIDVRDINTLEISDNOALPYDLRHRALLQHNQGVPIIAVONGICMSPFGFS 360  
DB 301 SAEPENIIDVRDINTLEISDNOALPYDLRHRALLQHNQGVPIIAVONGICMSPFGFS 360  
QY 361 MEMGLSIALIAVAVITTSFKFRPNOKA 387  
DB 361 MEMGLSIALIAVAVITTSFKFRPNOKA 387  
RESULT 2  
ID ABB76278 standard; Protein; 387 AA.  
XX ABB76278;  
AC  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Dirofilaria immitis cuticlin DiCut-1A.  
XX  
XX Cuticlin; helminth; parasite; DiCut-1A; vaccine; gene therapy;  
KW anthelmintic.  
XX  
OS Dirofilaria immitis.  
XX  
XX US2002037294-A1.  
PN  
XX 28-MAR-2002.  
PD  
XX 20-MAR-2001; 2001US-0812642.  
PF

XX  
PR 01-JUN-1999; 99US-0323427.  
XX  
XX (CHAN/) CHANDRASHEKAR R.  
PA (MORA/) MORALES T H.  
XX  
XX Chandrashekar R, Morales TH;  
PI  
DR MPI; 2002-401306/43.  
DR N-PSDB; ABL57372.  
XX  
XX  
PT New parasitic helminth cuticlin proteins and nucleic acids isolated  
PT form Dirofilaria immitis, useful for protecting animals from diseases  
PT caused by parasitic helminths or for inhibiting molting of filarid  
PT larvae in an animal -  
PS  
XX  
PS Claim 4: Page 21-22; 32pp; English.  
XX  
XX  
CC The present sequence is the protein sequence of novel  
CC Dirofilaria immitis cuticlin DiCut-1A, as predicted from an  
CC isolated cDNA clone (see ABL57372). The protein sequence has 91%  
CC identity to the Ascaris homologue. The invention provides D  
CC immitis and Brugia malayi parasitic helminth cuticlin proteins,  
CC nucleic acids, antibodies, compounds that inhibit cuticlin  
CC activity, and methods of obtaining them. Therapeutic compositions  
CC comprising such proteins, nucleic acids, antibodies and/or  
CC inhibitors, including genetic vaccines, recombinant virus vaccines  
CC and recombinant cell vaccines, are used to protect animals from  
CC diseases caused by parasitic helminths, and to inhibit the molting  
CC of filarid larvae in an animal.  
XX  
SO Sequence 387 AA:  
Query Match 100.0%; Score 2026; DB 23; Length 387;  
Best Local Similarity 100.0%; Pred. No. 7.1e-210;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MMRLIAFCTTLLALSYSIPVNGVGEPEIECGPTSTINFNTRNAFEGHYVVGGLYDQ 60  
DB 1 MMRLIAFCTTLLALSYSIPVNGVGEPEIECGPTSTINFNTRNAFEGHYVVGGLYDQ 60  
QY 61 ECGRNDEGGRQVAGISLPDSCNVARTRSINPRGIFVTTTVVISHPLFVTVDRAYRQ 120  
DB 61 ECGRNDEGGRQVAGISLPDSCNVARTRSINPRGIFVTTTVVISHPLFVTVDRAYRQ 120  
QY 121 CFYMEADKTVSAQIEVSETTAFQTQIYMPVCRYEILDGGPTGQVQFAIIGQPYVHKM 180  
DB 121 CFYMEADKTVSAQIEVSETTAFQTQIYMPVCRYEILDGGPTGQVQFAIIGQPYVHKM 180  
QY 181 TCDSEVDFFCVAVHSCFVDDNGDVTVEILNADGALDKYLLNNLEYPTDLMAAGEAHY 240  
DB 181 TCDSEVDFFCVAVHSCFVDDNGDVTVEILNADGALDKYLLNNLEYPTDLMAAGEAHY 240  
QY 241 KYADRSQLFYQCOISITTIKEPNSCEVRPQCSPOGFGAVKTGGAAPKAPAAAQRLKKR 300  
DB 241 KYADRSQLFYQCOISITTIKEPNSCEVRPQCSPOGFGAVKTGGAAPKAPAAAQRLKKR 300  
QY 301 SAEPENIIDVRDINTLEISDNOALPYDLRHRALLQHNQGVPIIAVONGICMSPFGFS 360  
DB 301 SAEPENIIDVRDINTLEISDNOALPYDLRHRALLQHNQGVPIIAVONGICMSPFGFS 360  
QY 361 MEMGLSIALIAVAVITTSFKFRPNOKA 387  
DB 361 MEMGLSIALIAVAVITTSFKFRPNOKA 387  
RESULT 3  
ID AAB84260 standard; Protein; 245 AA.  
XX AAB84260;  
AC  
XX  
DT 22-AUG-2001 (first entry)  
PF

```

XX      Amino acid sequence of a cuticlin polypeptide.
DE
XX
XX      Cuticlin: gene therapy; vaccine; helminth parasite.
KM
XX
OS      Brugia malayi.
XX
PN      US6248329-B1.
XX
PD      19-JUN-2001.
XX
XX      01-JUN-1999; 99US-0323427.
XX
XX      01-JUN-1998; 98US-0087435.
XX
PA      (CHAN/) CHANDRASHEKAR R.
PA      (MORA/) MORALES T H.
PI      Chandrashekar R, Morales TH;
XX
XX      WPI; 2001-396953/42.
DR      N-PSDB; AAB24839.
XX
XX      Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin
PT      protein is useful as a vaccine to prevent parasitic helminth infection
PT
XX
XX      Disclosure; Column 53-54; 29pp; English.
XX
XX      The specification describes a Dirofilaria immitis nucleic acid molecule,
CC      that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid
CC      molecule is useful as a probe to identify nucleic acid molecules, as a
CC      primer to produce nucleic acids, as a therapeutic reagent, e.g., gene
CC      therapy to inhibit cuticlin activity or production, or in a vaccine to
CC      prevent infection with helminth parasites. The cuticlin protein,
CC      antibodies raised against it, and inhibitory compounds of cuticlin may
CC      all be used in compositions to protect animals, especially mammals such
CC      as cats, dogs, and humans. The antibodies may be used to passively
CC      immunize an animal, or as reagents in assay to detect infection of
CC      helminths, or as tools to screen expression libraries to recover desired
CC      proteins. They may also be used to target cytotoxic agents to the
CC      parasite and kill it directly. The present sequence represents a cuticlin
CC      polypeptide.
XX
SQ      Sequence 245 AA;
Query Match 53.6%; Score 1086.5; DB 22; Length 245;
Best Local Similarity 80.6%; Pred. No. 1.3e-108;
Matches 195; Conservative 24; Mismatches 22; Indels 1; Gaps 1;
QY      3 IRLAFCITLIALSY-SIPVDNGVEGEPEICGPTSTINENTRNAFEGHYVVGLYDOE 61
DB      4 MOICSPFLSYMIASINMIPIDNGVESEPEICGPTSTIVNENTRNPEGHYAGLYSNQ 63
QY      62 GCRNDEGROVAGISLPEDSCNVAARTSLNRGIFVTTTVISFHPLEFVTVDRAYVQC 121
DB      64 DCRSDEGROVAGISLPEDSCNVAARTSLNRGIFVTVAVVITTHPOGITVDRTRYRQC 123
QY      122 FYMEADKTVAQIEVSEITTAFOQIYMPVPCRYEILDGPGTGPVOYANIGOPVYHKWT 181
DB      124 FYMEADKTVAQIEVSEITTVFAFQIYMPVPCRYEILDGPGTGPVOYANIGOPVYHKWT 183
QY      182 CDSETVDFCAVHSCFVDGNGDPTVELNADGALDKYLLNNLEYPTDLMAAGBAHYK 241
DB      184 CDSETVDFCALVHSCFVDGNGDSINLINNEGALDRYLLNNLEYPTDLMAAGBAHYK 243
QY      242 YA 243
DB      244 YA 245

```

```

ID      ABB76280 standard; Protein; 245 AA.
XX
XX      ABB76280;
AC
XX      12-AUG-2002 (first entry)
DT
XX
DE      Brugia malayi cuticlin BmCut-1A.
XX
XX      Cuticlin; helminth; parasite; BmCut-1A; vaccine; gene therapy;
KM      anthelmintic.
XX
XX      Brugia malayi.
OS
XX
XX      US2002037294-A1.
PN
XX      28-MAR-2002.
PD
XX      20-MAR-2001; 2001US-0812642.
XX
XX      01-JUN-1999; 99US-0323427.
XX
XX      (CHAN/) CHANDRASHEKAR R.
XX      (MORA/) MORALES T H.
PI      Chandrashekar R, Morales TH;
XX
XX      WPI; 2002-401306/43.
DR      N-PSDB; ABL57374.
XX
XX      New parasitic helminth cuticlin proteins and nucleic acids isolated
PT      from Dirofilaria immitis, useful for protecting animals from diseases
PT      caused by parasitic helminths or for inhibiting molting of filarial
PT      larvae in an animal.
XX
XX      Example 1; Page 28-29; 32pp; English.
XX
XX      The present sequence is the protein sequence of novel
CC      Brugia malayi cuticlin BmCut-1A, as predicted from an isolated
CC      partial cDNA clone (see ABL57374). The invention provides B.
CC      malayi and Dirofilaria immitis parasitic helminth cuticlin proteins,
CC      nucleic acids, antibodies, compounds that inhibit cuticlin
CC      activity, and methods of obtaining them. Therapeutic compositions
CC      comprising such proteins, nucleic acids, antibodies and/or
CC      inhibitors, including genetic vaccines, recombinant virus vaccines
CC      and recombinant cell vaccines, are used to protect animals from
CC      diseases caused by parasitic helminths, and to inhibit the molting
CC      of filarial larvae in an animal.
XX
SQ      Sequence 245 AA;
Query Match 53.6%; Score 1086.5; DB 23; Length 245;
Best Local Similarity 80.6%; Pred. No. 1.3e-108;
Matches 195; Conservative 24; Mismatches 22; Indels 1; Gaps 1;
QY      3 IRLAFCITLIALSY-SIPVDNGVEGEPEICGPTSTINENTRNAFEGHYVVGLYDOE 61
DB      4 MOICSPFLSYMIASINMIPIDNGVESEPEICGPTSTIVNENTRNPEGHYAGLYSNQ 63
QY      62 GCRNDEGROVAGISLPEDSCNVAARTSLNRGIFVTTTVISFHPLEFVTVDRAYVQC 121
DB      64 DCRSDEGROVAGISLPEDSCNVAARTSLNRGIFVTVAVVITTHPOGITVDRTRYRQC 123
QY      122 FYMEADKTVAQIEVSEITTAFOQIYMPVPCRYEILDGPGTGPVOYANIGOPVYHKWT 181
DB      124 FYMEADKTVAQIEVSEITTVFAFQIYMPVPCRYEILDGPGTGPVOYANIGOPVYHKWT 183
QY      182 CDSETVDFCAVHSCFVDGNGDPTVELNADGALDKYLLNNLEYPTDLMAAGBAHYK 241
DB      184 CDSETVDFCALVHSCFVDGNGDSINLINNEGALDRYLLNNLEYPTDLMAAGBAHYK 243
QY      242 YA 243
DB      244 YA 245

```

	RESULT	5
ID	AAB84259 standard; Protein; 271 AA.	
XX	AAB84259	
XX	AAB84259;	
XX		
DT	22-AUG-2001 (first entry)	
XX		
DE	Amino acid sequence of a cuticlin polypeptide.	
XX		
KW	Cuticlin; gene therapy; vaccine; helminth parasite.	
XX		
OS	Dirofilaria immitis.	
PX	US6248329-B1.	
PN		
PD	19-JUN-2001.	
XX		
PF	01-JUN-1999; 99US-0323427.	
PR	01-JUN-1998; 98US-0087435.	
XX		
PA	(CHAN/) CHANDRASEKHAR R.	
PI	(MORA/) MORALES T H.	
PI	Chandrasekhar R, Morales TH;	
DR	WPI; 2001-396953/42.	
N-PSDB:	AAH24830, AAH24832.	
XX		
PT	Novel nucleic acid molecule encoding a Dirofilaria immitis cuticlin	
PT	protein is useful as a vaccine to prevent parasitic helminth infection	
XX		
Example 1; Column 47-50; 29pp; English.		
XX		
PS	The specification describes a Dirofilaria immitis nucleic acid molecule,	
CC	that encodes a cuticlin protein. The Dirofilaria immitis nucleic acid	
CC	molecule is useful as a probe to identify nucleic acid molecules, as a	
CC	primer to produce nucleic acids, as a therapeutic reagent, e.g., gene	
CC	therapy to inhibit cuticlin activity or production, or in a vaccine to	
CC	prevent infection with helminth parasites. The cuticlin protein,	
CC	antibodies raised against it, and inhibitory compounds of cuticlin may	
CC	all be used in compositions to protect animals, especially mammals such	
CC	as cats, dogs, and humans. The antibodies may be used to passively	
CC	immunize an animal, or as reagents in assay to detect infection of	
CC	helminths, or as tools to screen expression libraries to recover desired	
CC	proteins. They may also be used to target cytotoxic agents to the	
CC	parasite and kill it directly. The present sequence represents a cuticlin	
XX	polypeptide.	
XX		
SO	Sequence 271 AA;	
Query Match	38.3%; Score 776.5; DB 22; Length 271;	
Best Local Similarity	72.4%; Pred.No. 4.8e-75;	
Matches 144; Conservative 24; Mismatches 22; Indels 9; Gaps 2;		
OY	134 IEVSITTAPOQTQIVPMVCRREIILDGPGTGPVFATIGCVYHKWTCDSETVDTFCAV	193
DB	15 LEVSEMTTAFOTQVPMPVCREILEGGPTGAPVRFMAGDVHYHKWTCDSSTTTFCAL	74
OY	194 VHSCEVDGDNDJVEIILNADSCALDKYLLNNNEYPTDLMAGOEAHVYKYADRSQLFYOCQ	253
DB	75 VHSVCVDDGKRDVAEILNEBGCALDKYLNNLEIFTDLDMAGEAHHVKRADNSSELYOCQ	134
OY	254 ISITIKPENSECVRPQCSEPGFGAVKTG-----GAAKPAAAAOL--RLTKRSASP	304
DB	135 ISITIKPEHSHCPRTPEPGFGLKSGQGFAAYKSAAPRAPASLSLSPRIKKRSINS	194
OY	305 ENIIDVPRDIITLTLSIDSN 323	
	: : :       : : :   : : :   : : :	

Db	195	DNTVDVSTGFSYVDITEEN	213	
		RESULT 6		
		ABB76279		
XX	ID	ABB76279	standard; Protein; 271 AA.	
XX	AC	ABB76279;		
XX	DT	12-AUG-2002	(first entry)	
XX	DE	Dirofilaria immitis	cuticlin DiCut-1B.	
XX	KW	Cuticlin; helminth; parasite; DiCut-1B; vaccine; gene therapy; antelmintic.		
XX	OS	Dirofilaria immitis.		
XX	PN	US2002037294-A1.		
XX	PD	28-MAR-2002.		
XX	PF	20-MAR-2001; 2001US-0812642.		
XX	PR	01-JUN-1999; 99US-0323427.		
XX	PA	(CHAN/) CHANDRASHEKAR R.		
XX	PI	(MORA/) MORALES T H.		
XX	PI	Chandrashekar R, Morales TH;		
XX	DR	WPI: 2002-403306/43.		
XX	DR	N-PSDB; ABL57373.		
PT	PT	New parasitic helminth cuticlin proteins and nucleic acids isolated		
PT	PT	form <i>Dirofilaria immitis</i> , useful for protecting animals from diseases		
PT	PT	caused by parasitic helminths or for inhibiting molting of filarial		
PT	PT	larvae in an animal -		
XX	XX	Claim 4; Page 25-26; 32pp; English.		
PS	PS	The present sequence is the protein sequence of novel		
CC	CC	<i>Dirofilaria immitis</i> cuticlin DiCut-1B, as predicted from an		
CC	CC	isolated cDNA clone (see ABL57373). The protein sequence has 81%		
CC	CC	identity to the <i>Ascaris</i> homologue. The invention provides D.		
CC	CC	<i>immitis</i> and <i>Brugia malayi</i> parasitic helminth cuticlin proteins,		
CC	CC	nucleic acids, antibodies, compounds that inhibit cuticlin		
CC	CC	activity, and methods of obtaining them. Therapeutic compositions		
CC	CC	comprising such proteins, nucleic acids, antibodies and/or		
CC	CC	inhibitors, including genetic vaccines, recombinant virus vaccines		
CC	CC	and recombinant cell vaccines, are used to protect animals from		
CC	CC	diseases caused by parasitic helminths, and to inhibit the moulting		
CC	CC	of filarial larvae in an animal.		
XX	XX			
XX	Sequence	271 AA;		
XX	Query Match	38.3%; Score 776.5; DB 23; Length 271;		
XX	Best Local Similarity	72.4%; Pred. No. 4.8e-75;		
XX	Matches 144; Conservative	24; Mismatches 22; Indels 9; Gaps		
XX	XX			
QY	134	LEVEITTAFTQTVPMVVCYREILIDGPTQPOFAIIGQPVYHKMTCDSETVTFCAV	193	
Db	15	LEVEEMTAFQTQVPMVPCVRCYREILEGGTAPVRFAMIGYHVHKMTCDSETDTPFCAL	74	
QY	194	VHSEVDDGNDQVEIINADGALDKYLLNNLEVPDLMAGOEAHVYKVAARSOQFYCOQ	253	
Db	75	VHSCVVDGKDAVEILNEBSCALDKYLLNNLEYITDLMAGOEAHVYKVAARSELVYCOQ	134	
QY	254	ISITIKPEPNSCVAPQCSPEPGFAGVTKG-----GAAKAPAAAOI---BLTKRSAP	304	
Db	135	ISITIKPEHSCRPQCTEPQGFCAINSGGSGFAAVKSAAPASLSPLIKRSINS	194	
QY	305	ENIDVVRTDINTLEISDNN	323	

Db 195 DNTVDSTGFSTVDITEEN 213

RESULT 7  
ID ABB59975 standard; Protein; 462 AA.

XX ABB59975;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 6717.

KM Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

PI N-PSDB; ABL04078.

DR MPI: 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

PS Disclosure; SEQ ID NO 6717; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
sequences (AB101840-AB16175) and the encoded proteins  
(ABBS7737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 462 AA;

Query Match 9.8%; Score 199; DB 22; Length 462;

Best Local Similarity 22.2%; Pred. No. 3e-12; Mismatches 150; Indels 124; Gaps 20;

Matches 94; Conservative 55; Mismatches 150; Indels 124; Gaps 20;

QY 25 VEGEPPEIEC--GPTSTI-----NENTRAPEGHVYVKGILYDDEGCRNDEGGROVAG--- 74

DB 59 VEPSVRKICLGSMLITIKDAPNHER-GIFSGLMYPKGLSKNSTCLSEV--RDHVSRLR 115

QY 75 ISLPDSCNVARIRSLNPRGIFVTYTVVISFHPLEFVKVRAVRVOCFYMEADKTV---- 130

DB 116 YKLPLNSCN-TMRETDGSGIEFPNTVLDPLKLTLDLGRGHVRCAYSRDAAMPKK 174

QY 131 -----SAQIEVSEITTAFOQTIVPM 150

DB 175 YLRKHAKQPAFRSDRREYGRSLDKQDDDLDEEDYDANAPQEDDV-----NNIIPM 230

QY 151 PVC-----RYELDGGPRPGQVQFAITGQPVYHKWTCSEYVDFTCAYVHSCFVDNGN 203

Db 231 PGCHMKIYNDEHKIADVDKIGDPLTIYI-----SIDKQKV-YGLHVTDCIVRDGL 279

QY 204 G-DTVEILNDGSCALDKYTLNNLEYPIDIMAGQ--EAHYVKYADRSQLFOCOISITIK 259

Db 280 GWGQRLVGEDGCPMDIEIKQGFNTODRLANVTFFAHKFPYT--TSYIQCWVRICAL 337

QY 260 EPNSECVRPQCSSEPQGFAGVKTGGAARPAAPAOIRLLKKRSAPENI--IDVRTDINT 316

Db 338 EDPTCGEAPQCS-----GKRPKRQAAD-----SKREDGLPATIEVFSGIYV 379

QY 317 LEISDDNOLPVDLRHALLQHNQCPVITLAAVONGICSPGSGSMFG-----LSIALIA 371

Db 380 NENENAN-----DSDEDAVYKEK-----TLDALCVSQRFFAIAIAGLIIMAVVA 427

QY 372 AVI 374

Db 428 AVL 430

RESULT 8  
ID ABB70061 standard; Protein; 604 AA.

XX ABB70061;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 36975.

KM Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

PI N-PSDB; ABL14164.

DR MPI: 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

PS Disclosure; SEQ ID NO 36975; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
sequences (ABBS7737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 604 AA;

Query Match 8.2%; Score 166; DB 22; Length 604;

Best Local Similarity 22.0%; Pred. No. 1.7e-08; Mismatches 135; Indels 66; Gaps 17;

[illegible]

CC	Result 9
CC	ABB67889
XX	ID ABB67889 standard; Protein; 1638 AA.
XX	AC ABB67889;
XX	DT 26-MAR-2002 (first entry)
XX	DE Drosophila melanogaster polypeptide SEQ ID NO 30459.
XX	KW Drosophila: developmental biology; cell signalling; insecticide;
XX	KW pharmaceutical.
XX	OS Drosophila melanogaster.
XX	PN WO200171042-A2.
XX	PD 27-SEP-2001.
XX	PF 23-MAR-2001; 2001WO-US09231.
XX	PR 23-MAR-2000; 2000US-191637P.
XX	PR 11-JUL-2000; 2000US-0614150.
XX	PA (PEKE ) PE CORP NY.
XX	PI Venter JC, Adams M, Li PWD, Myers EM;
XX	PI WPI: 2001-656860/75.
DR	N-PSDB; ABLI1992.
XX	PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX	PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX	PT interactions -
XX	PS Disclosure; SEQ ID NO 30459; 21pp + Sequence Listing; English.
XX	XX
XX	XX
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC	sequences (AB101840-AB16175) and the encoded proteins
CC	(AAB57737-AB12072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 1638 AA;

Query Match	8.0%;	Score 163;	DB 22;	length 1638;
Best Local Similarity	19.8%;	Pred. No. 1.8e-07;		
Matches	85;	Conservative	62;	Mismatches 155;
			Indels	128;
			Gaps	16;

QY	27	GEPEIEGPGPSIYIN-----ENTNABEGHYVYVGLYDQESCRNDBGROVAGISLP-78
Db	1060	GRSGEGECMDYQCNBEGMEFTIRTEBGLGRITYYGFDYDRCFFRGNGTAVNLRISGPQ 1119
QY	79	-FDSCNARTRSINPRGIEVYTTVVISFHLPEFTKVDARAYRQCF-----122
Db	1120	GYPCCGTQR-----XGDLTLNIVVQGFSDNNQTSMDKRYNLTLCLFRBGEAVYSSGIY 1173
QY	123	-----YMEADTVSAQIIVSEITTAFTOTQIVPMPCRYEILDGPT-----GOV 167
Db	1174	AGSGSPIPIEYLRPAENLTSSKVLRS-----ILYQGRPTTTAVGDPL 1215
QY	168	QFALIGQPVYHKKTQCESEYVDFECAYVHSCFVDGNGDTVEILMDGALDKYLLNNLEY 227
Db	1216	TFRLEADGGINHYT-----DIFATNVVA--KDPISGRSIOQLIDRFGCPVDFEYFPELDK 1267
QY	228	PTD---LMAGQEAHVYKVADRSOLFYOQISITITIKEPNSECVRPOCSBPGFCAVKTGA 284
Db	1268	LRGBDTLEA--RENAKIKIPESNLFVEA-----TVRSCRGCPAPVCPGAPG-----1312
QY	285	AAMPAAAOURLKKNS--AEPE-----NIDVRDINMLEISDNOALPVD--329
Db	1313	KQPSFGRRRRSINTTETPEPELALBEGSSQLEASTLDEVTIVNSVTSAITGOVPLNET 1372
QY	330	-----LRRALLQNGQPVITLAAYONGIOMCSBFGSMFGLSIA 368
Db	1373	OLGEBKTRETEEPQVREMIIEVEFETREIEKESYPRKLVAPVETVQMP---AEYHGLITA 1429
QY	369	LIAAVITITS 378
Db	1430	IIILMILIFS 1439

RESULT	10	
AB67239		
ID	AB67239	standard; Protein; 611 AA.
XX		
AC	AB67239:	
XX		
DT	26-MAR-2002	(first entry)
XX		
DE	Drosophila melanogaster	polypeptide seq ID NO 28509.
XX		
KW	Drosophila: developmental	biology; cell signalling; insecticide;
KW	pharmaceutical.	
XX		
OS	Drosophila melanogaster.	
XX		
PN	WO200171042-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	23-MAR-2001;	2001WO-US09231.
XX		
PR	23-MAR-2000;	2000US-191637P.
PR	11-JUL-2000;	2000US-0614150.
XX		
PA	(PEKE )	PE CORP NY.
XX		
PI	Venter JC, Adams M,	LA PWD, Myers EW;
XX		
WI	WI: 2001-656860/75.	
DR	N-PADB: ABL11342.	
XX		
XT	New isolated nucleic acid detection reagent for detecting 1000 or more	

PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX  
PS Disclosure; SEQ ID NO 28509; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (ABBS7737-ABBS72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 611 AA;  
SQ  
  
Query Match 8.0%; Score 162.5; DB 22; Length 611;  
Best Local Similarity 21.4%; Pred. No. 4.1e-08;  
Matches 63; Conservative 51; Mismatches 126; Indels 55; Gaps 13;  
  
QY 30 EIECGPTSTINFTNNAFEGHYVVKGLYDQEGCRNDGGRQVAGISLPD---SCNVA 85  
DB 88 QVOCCKTHMRVNIIEFDRPFYGMIFSKGYSDPHCVHLKPG--TGHLSATFEIPLNSG 145  
QY 86 RTSLN-----PRGIFVTVTVISFHPLEFVKVRAVAVOC---FVMEADKTVSAQ 133  
DB 146 SSANHNAGYAGPAPSSGYVENITIIQYDPYQEVWQARKLRCTWDFYEKAVTFRRPQ 205  
QY 134 IEVSEITTA-----FQTQIVMPVCRYEILDG-GPTGQPVQFAI-IGQPYVHKWTC 185  
DB 206 VDMHLAVTANFLGDNLQ-----CWMQIQVKGKPMASEVSGIVKIGQFTWTVLAIK 257  
QY 186 TVDFECAVVHSCFVDDNGDVTVELNADGALDKYLLNLEYPDLMAGQEAHY----- 239  
DB 258 E-NKFEDLVNRCVAHDKRAPIQLVDQNGCVVRPKIMSKFOKIKNF--GPSASVVSFA 314  
QY 240 --KKYADRSLQFYQCOISITIKENSCVPRQCEPQGFAGVAKTGAAPRAAA 292  
DB 315 QAEKFPDSMNVHFCVQIVC---RYNCPKPKC-----GGLPGEGELPQIGA 359  
  
RESULT 11  
ABB70054  
ID ABB70054 standard; Protein; 611 AA.  
XX  
AC ABB70054;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 36954.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
XX  
XX WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.

DR N-PSDB; AB114157.  
XX  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX  
PS Disclosure; SEQ ID NO 36954; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (ABBS7737-ABBS72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 611 AA;  
SQ  
  
Query Match 8.0%; Score 162.5; DB 22; Length 611;  
Best Local Similarity 21.4%; Pred. No. 4.1e-08;  
Matches 63; Conservative 51; Mismatches 126; Indels 55; Gaps 13;  
  
QY 30 EIECGPTSTINFTNNAFEGHYVVKGLYDQEGCRNDGGRQVAGISLPD---SCNVA 85  
DB 88 QVOCCKTHMRVNIIEFDRPFYGMIFSKGYSDPHCVHLKPG--TGHLSATFEIPLNSG 145  
QY 86 RTSLN-----PRGIFVTVTVISFHPLEFVKVRAVAVOC---FVMEADKTVSAQ 133  
DB 146 SSANHNAGYAGPAPSSGYVENITIIQYDPYQEVWQARKLRCTWDFYEKAVTFRRPQ 205  
QY 134 IEVSEITTA-----FQTQIVMPVCRYEILDG-GPTGQPVQFAI-IGQPYVHKWTC 185  
DB 206 VDMHLAVTANFLGDNLQ-----CWMQIQVKGKPMASEVSGIVKIGQFTWTVLAIK 257  
QY 186 TVDFECAVVHSCFVDDNGDVTVELNADGALDKYLLNLEYPDLMAGQEAHY----- 239  
DB 258 E-NKFEDLVNRCVAHDKRAPIQLVDQNGCVVRPKIMSKFOKIKNF--GPSASVVSFA 314  
QY 240 --KKYADRSLQFYQCOISITIKENSCVPRQCEPQGFAGVAKTGAAPRAAA 292  
DB 315 QAEKFPDSMNVHFCVQIVC---RYNCPKPKC-----GGLPGEGELPQIGA 359  
  
RESULT 12  
ABBS8337  
ID ABBS8337 standard; Protein; 758 AA.  
XX  
AC ABBS8337;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 1803.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
XX  
XX WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
DR



PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX MPI; 2001-656860/75.  
 DR N-PSDB; ABL02440.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 1803; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB157737-AB172072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 758 AA:  
 SQ  
 Query Match 7.4%; Score 150; DB 22; Length 758;  
 Best Local Similarity 23.6%; Pred. No. 1.3e-06;  
 Matches 81; Conservative 50; Mismatches 114; Indels 98; Gaps 20;  
 OY 31 IECGPTSIINFTNRAFGHYVVGKLYDQECGRNDEGGRQVAGISLPD-----SC 82  
 DB 393 ICRSGEMITKIRTSKLFQGYAKGA--PKCAVNVN-----SLEDFMKRYNDELC 444  
 OY 83 NVART---KSLNPRGIFVTYVVISFHPLEFVKVDRAYVOCFYMEADKTVSAQIEV--- 136  
 DB 445 NVQSAVYGYVM-----DIVIGHDMIVYSSDLGLAVSCQYDLTKNTVYANNVDLGYT 496  
 OY 137 SETTAFQTOIY---PMPCRYEILDGPTGQPVQFAIGQVYHKMTCDSFVDT--F 190  
 DB 497 GEIESLSLEIITIDSPNVIMKTITARDG---SDMKRIAEVGDPLALRF---ETVDANSPY 549  
 OY 191 CAVHSCFVDGNGDPTVEI--LNADGALDKYLLNNLEYPTDLMAGQEAHVYKKA--DRQ 247  
 DB 550 ELFVRLVAMDGT-DSAEITLIDANGCPTDQYIMSAMQ-----KLANNRKV 594  
 OY 248 LFYQCOISITIKPEPSECVR-----PQSEF-----QGFQAVK----- 280  
 DB 595 LLSQED---AFKFPSESELVQFALVTPCIPRC-EPYICONDENGELKSLISGRKRKRYL 650  
 OY 281 --TGGAARPAALQRLKRSAPENIIDVRTDINTLEISD 321  
 DB 651 NGTDGVELAIKSEROKRDVSHQAGDENILVQ---SIQIDND 689  
 RESULT 13  
 ABB58185  
 ID ABB58185 standard; Protein; 833 AA.  
 XX  
 AC ABB58185;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 1347.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN MO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX

PE 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX MPI; 2001-656860/75.  
 DR N-PSDB; ABL02288.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 1347; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB157737-AB172072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 833 AA:  
 SQ  
 Query Match 7.0%; Score 142; DB 22; Length 833;  
 Best Local Similarity 22.7%; Pred. No. 1.1e-05;  
 Matches 88; Conservative 51; Mismatches 128; Indels 120; Gaps 22;  
 OY 31 IECGPTSIINFTNRAFGHYVVGKLYDQECGRNDEGGRQVAGISLPD-----SC 82  
 DB 381 ICRSGEMITKIRTSKLFQGYAKGA--PKCAVNVN-----SLEDFMRYNDELC 432  
 OY 83 NVART---KSLNPRGIFVTYVVISFHPLEFVKVDRAYVOCFYMEADKTVSAQIEV--- 136  
 DB 433 NVQSAVYGYVM-----DIVIGHDMIVYSSDLGLAVSCQYDLTKNTVYANNVDLGYT 484  
 OY 137 SETTAFQTOIY---PMPCRYEILDGPTGQPVQFAIGQVYHKMTCDSFVDTFCAY 193  
 DB 485 GEIESLSLEIITIDSPNVIMKTITARDG---SDMKRIAEVGDPLALRFEL-VEPNSPYELF 540  
 OY 194 VHSCEFVDGNGDPTVEI--LNADGALDKYLLNNLEYPTDLMAGQEAHVYKKA--DRSOLF 250  
 DB 541 VRELVAMDGS-DSAEITLIDANGCPTDQYIMGTIO-----KLANNRKVL 585  
 OY 251 QCOISITIKPEPSECVR-----PQSEFQGFQAVKGTGAARPAALQRL----- 297  
 DB 586 QFD---AFKFPSESELVQFALVTPCIPRC-EPV-----ICSEEDASGELKSLISYG 633  
 OY 298 -KKRSAPENIIDVRTDINTLEISDNOALPVDLRHALQHNQPVILAAVQNGICMSP 356  
 DB 634 RKRKRY-----LN-----GTDGAELISTYHR-----RDQGPV----- 661  
 OY 357 GFSGMEGLSIALAAVITTSFRRP 383  
 DB 662 -----DNNILMQSIQITQKFGQP 681  
 RESULT 14  
 ABB70535  
 ID ABB70535 standard; Protein; 692 AA.  
 XX  
 AC ABB70535;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX

Query Match	Best Local Similarity	Matches	6.9%	Score 140;	DB 22;	Length 692;
31	IECGSTTITNFMNNAFEHGYVKGIVDOEGCR----	NDEGGRQVAGISLFPDSCNVAR	86			
358	VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--	LDITMGAGDCN----	410			
87	TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS----	AOIEVSEITTAFL	143			
411	TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN	467				
144	QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN	203				
468	SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--	PYGIFASSCVAMAKD	521			
204	GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER	261				
522	ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----	574				
262	NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD	321				
575	---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD	626				
322	DNQ	324				
627	EKR	629				

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMGAGDCN----- 410

87 TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS-----AOIEVSEITTAFL 143

411 TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN 467

144 QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN 203

468 SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--PYGIFASSCVAMAKD 521

204 GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER 261

522 ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----- 574

262 NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD 321

575 ---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD 626

322 DNQ 324

627 EKR 629

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMGAGDCN----- 410

87 TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS-----AOIEVSEITTAFL 143

411 TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN 467

144 QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN 203

468 SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--PYGIFASSCVAMAKD 521

204 GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER 261

522 ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----- 574

262 NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD 321

575 ---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD 626

322 DNQ 324

627 EKR 629

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMGAGDCN----- 410

87 TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS-----AOIEVSEITTAFL 143

411 TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN 467

144 QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN 203

468 SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--PYGIFASSCVAMAKD 521

204 GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER 261

522 ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----- 574

262 NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD 321

575 ---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD 626

322 DNQ 324

627 EKR 629

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMGAGDCN----- 410

87 TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS-----AOIEVSEITTAFL 143

411 TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN 467

144 QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN 203

468 SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--PYGIFASSCVAMAKD 521

204 GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER 261

522 ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----- 574

262 NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD 321

575 ---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD 626

322 DNQ 324

627 EKR 629

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMGAGDCN----- 410

87 TRSLNPGCIFVYTVTVISFHPLEFVKVDRAYVQCFYMEADKTVS-----AOIEVSEITTAFL 143

411 TQSVY--GVY--SNIVVLQHSVVMKADKIYVVKCTYDMSKNITFGMPRINDPEKHIN 467

144 QTQIVMPVCRYEIIDGPTGQVFAITGCPVYHKWCTDSETVDFCAVHSCFEVDGN 203

468 SSPEAPPP--RIRIID--TROREVETVIRIGRLNFRIEPEDT--PYGIFASSCVAMAKD 521

204 GDT--VEINADGCAIDKLLNLLEFPTDLMAQGEAH--YKYVDRSQOLFQOCIOISTIKER 261

522 ARTSKRIIDDDGCPDPPIFPG--PTAAGNALQSTYEAFREFSESQVIFQCNVXY----- 574

262 NSECRPPCCPEQFGAVKGTGAALKPAAALQRLTLKKRSPENINIDVPTDINTLEISD 321

575 ---CLGFP--CEP---AVCEMNMDSFESLGRRRRRSIESNDIKSEDDMNISQELIVDFGD 626

322 DNQ 324

627 EKR 629

Sequence 692 AA;

Query Match

Best Local Similarity 20.8%; Pred. No. 1.4e-05;

Matches 63; Conservative 59; Mismatches 141; Indels 40; Gaps 16;

31 IECGSTTITNFMNNAFEHGYVKGIVDOEGCR-----NDEGGRQVAGISLFPDSCNVAR 86

358 VHCKDTRIAVOVRINKKPNRGRTYALG--RSEFCNIDIVISDFR--LDITMG

ID	ABBS8426	ABBS8426 standard; Protein; 744 AA.
XX	AC	ABBS8426;
XX	DP	26-MAR-2002 (first entry)
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 2070.
XX	KW	Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.
XX	OS	Drosophila melanogaster.
XX	PN	WO200171042-A2.
XX	PD	27-SEP-2001.
XX	PE	23-MAR-2001; 2001WO-US09231.
XX	PR	23-MAR-2000; 2000US-191637P.
XX	PR	11-JUL-2000; 2000US-0614150.
XX	PA	(PEKE ) PE CORP NY.
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	DR	WPI; 2001-656860/75.
XX	DR	N-PSDB; ABL02529.
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX	PS	Disclosure; SEQ ID NO 2070; 21pp + Sequence Listing; English.
XX	CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL01840-ABL016175) and the encoded proteins (ABBS7737-ABBS72072).
XX	CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	SO	Sequence 744 AA;
XX	Query Match	6.7%; Score 135.5; DB 22; Length 744;
XX	Best Local Similarity	20.5%; Pred. NO. 4.7e-05;
XX	Matches	82; Conservative 63; Mismatches 142; Indels 113; Gaps 20;
QY	1	IECGFTSTINENTRNAEFGHYVKGLYDQCGRNDEGGROYAGISLPFD-----SC 82
DB	380	IECGGDMILARIRTSKLENGKYAKG--SPKSCSYD-----VKSALDELHMNYHDLRC 431
QY	83	NVARRSRLNPKRIEFTTVTVISFHFLEFYTKVDRAIRVQCFYMEADKTVSAQIEV---SEI 139
DB	432	NVRQSTA---GRYV-NDIIIDHMIYSSDLGIALAQYDLTKNSVSGVDLDRGDI 486
QY	140	TTAFCTQIV---PMVPCREILDDGGPTGQPOFATIGQGVYHKMNCDSFTVTPCAVYHS 196
DB	487	MPALSEEVIVESPNIYMRITSRDG---SDMMSAEYGDPLAKFELVDEQSPYELFIREL 543
QY	197	CFVDDNGDFTVEILNADGALD-----KYLNNMLEYPTDLMAGOEAAHVYK 242
DB	544	VAMQDVDSSEITLIDNSNGCPTHFIMGPIRYKSVSGKMILSNFD-----AKFV 591
QY	243	ADRSOLFVQCQISTITKEPNSECVPRQCEPQGFQAVKGTGAAGAPAAQAQLRL---K 298
DB	592	P--SSEVVOFRLVYPCMPSCPEV---QCEQEDTSGEFRS-----LTSYGRK 633

[illegible]

RESULT 16	
ABB64380	
ID	ABB64380 standard; Protein; 638 AA
XX	

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB1161716-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB57757-AB572072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

Query Match	6.48;	Score 130.5;	DB 22;	Length 638;
Best Local Similarity	21.78;	Pred. No. 0.00013;		
Matches	55;	Conservative 43;	Mismatches 110;	Indels 45;
				Gaps 12;

```

QY 141 TAFQTQIYPMVPCRYEYELIDG-GPTGQGVQFAIIGQPVYHKMTCDSSEVDFECV----- 193
Db 124 --FPGDNVD--CMETIQHGKGPAPVPVSGIV--PL-----GSTLLVLVAINDYRGE 168
QY 194 ----VHSCFVDDGNGDPVEIILNADGALDRLYLLNLT-----EYPTDJMAGQEAHYKYA 243
Db 169 FDMMYKSCVASDVGSHYINLSDERGCYLRPRKMSIRFLKARAPDERATVITYAFHAFKFP 228
QY 244 DRSLFYQCOISI 256
Db 229 DALSVHAKCKVEI 241

```

RESULT 17  
ABB64367  
ID ABB64367 standard; Protein; 699 AA.  
yy

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB161716-AB130511), expressed DNA sequences (AB101840-AB161715) and the encoded proteins (AB57737-ABR2072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Query Match	6.18;	Score 124.5;	DB 22;	length 699;
Best Local Similarity	18.78;	Pred. No. 0.00066;		
Matches	80;	Conservative	67;	Mismatches 182;
				Indels 99;
				Gaps 19;

```

Db 312 TDNYGQGYGHEAGSTGAGTYFENITIVIOYDPOVEWMOQARKLCTMHWQYKESVTER-- 369
QY 136 VSEITTAFOQIVMPV-----CRVELLOS-GPTGQPVQFAI-IGQPVY 177
Db 370 -----PPVMDLDVBRADPAGDNVGCWMOIQVKGFWASEVSLVXIGQTM 416
QY 178 HKWTCDSEYVDTFCVAVVHSCFVDGNGDVEILNADGALDKYLLN-----NLEYPTDL 231
Db 417 MVLAIKDD--SKFMDLVKNCVAHDGKRAPILQVDRGCVTPKMLSRRTKIKNGASASV 475
QY 232 MAGGEAHYKADRSQLEFYQCQISITIKERNSEC-----VRPOCS---EPQ- 274
Db 476 LSYAHFOAFKFRPDMSEVHFQCTIQICRYHCEPCCSAETNLODVHNLQVGPESQYGPPO 535
QY 275 GFGAVTKGGAAKPAAAOQLRLKKRSAPR-----NIIDVTDINTLEISDNOA----- 325
Db 536 HVDVAVHASAIGKRDRERRVRRARAFAEPQVGNRLIKV-VSSGDLTFALIDDOAGNGS 594
QY 326 -----LPVDLRHRALLQHNQPVILAAVONG-ICMSPFGFSMFGSLIALIAVILR-- 376
Db 595 STNGANGVD-----RSPQTMVFPRLREGLICMTTPGFAITLIVLGLITVLSCLTSA 646
QY 377 -ISFKRP 383
Db 647 VLYVRLRP 654

RESULT 18
ABB68164
ID ABB68164 standard; Protein; 418 AA.
XX AC ABB68164;
XX DF 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 31284.
XX KM Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI: 2001-656860/75.
XX DR N-PSDB; ABL12267.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 31284; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX CC sequences (AB101840-AB116175) and the encoded proteins
XX CC (ABB57737-ABB72072).

```

```

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 418 AA:
Query Match 5.9%; Score 119.5; DB 22; Length 418;
Best Local Similarity 19.8%; Pred. No. 0.001;
Matches 78; Conservative 67; Mismatches 103; Indels 65; Gaps 16;
QY 27 GEPEECPTITINE-----NTRNAFEGHYVYKGL-YDQEGCRND-EGGRQVAGISL- 77
Db 29 GSHVYHCEDDMRVDIGLPDAESKQDAPQIYLEGLKVPDERCOPQIDGSLAVFRLSL 88
QY 78 PEDSCNVARTRSLNPRGIFVTTVVISHPHLF-----YTKVBRAYVQCFY 123
Db 89 DFYECGV--TRMVO-----ITGKKVYHKIIESTSKEIYVSKCITTASPAIVMMNA 141
QY 124 MEADKTVSA-----QIEVSEITTAFOQIVMPVPCRYEILDGPTG 164
Db 142 TTGSSSTSSGGIHGLYKRDVLPAGRPEDDETTTSL-TKRAPEPLSTIGVSQDGKF 200
QY 165 QPVQFAITGQPVYHKWTCDSFTVDFCAVHSCFVDGNGDVEILNADGALDKYLLN 224
Db 201 TRDLTVKSGTFLTMEINDEDSAPYGLGVNLDVTDHTSS-ETLIFKGVDPDYLPEN 259
QY 225 LEYPTDLMAQGEAHYKADRSQLEFYQCQISITIKERNSECVRPOCSFPQ-GFGAVKTG 283
Db 260 FNTIDGDLTSAKFAKFPDSSYVOFRATVAVCL---DKCLGQCSSNNQVGFGRKREI 315
QY 284 AAAPRAAAOQLRL-LKRSAPR-----IIVRTDINTLEISDNOQLPVDLRHRALLQH- 338
Db 316 SSANKVYEISLAMPLOQVDIGVKNKEVLOLEKRIELKLA-NQRLARNSRGNFAMQ 373
QY 339 --NGQPVILAAVONGICMSPFGFSMFGSLIAL 369
Db 374 PASAQPAFVVDRELGHLISAGSGAASNGLSIAL 406

RESULT 19
ABB71434
ID ABB71434 standard; Protein; 2284 AA.
XX AC ABB71434;
XX DF 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 41094.
XX KM Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI: 2001-656860/75.
XX DR N-PSDB; ABL15537.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -

```



FT	/product= PHL
XX	
PN	EP65404-A.
XX	
PD	08-JAN-1992.
XX	
PE	27-MAY-1991; 91EP-0810399.
XX	
PR	30-NOV-1990; 90JP-0329911.
PR	01-JUN-1990; 90JP-0141678.
XX	10-AUG-1990; 90JP-0210535.
PA	(CIBA ) CIBA GEIGY AG.
PI	Iwasaki Y, Shimo H, Suzuki K, Ghisalba D, Nishikawa Y;
PI	Kawahara T, Kangawa K;
XX	
DR	WPI: 1992-010570/02.
N-PSDB:	AAQ20269.
PT	Novel DNA encoding peptidyl hydroxyglycine-N-C lyase (PHL) -
PT	used to prepare PHL which can be used in the amidation of
XX	peptide(s) e.g. human calcitonin.
CC	
PS	Claim 4; Page 18; 28pp; English.
XX	
CC	The sequence was deduced from a cDNA insert from PAE-III-202-4
CC	(FERM BP-3172). The vector serves as a source for a DNA fragment
CC	encoding PHL for the construction of an expression vector for the
CC	prep'n. of recombinant PHL. The protein may be truncated to
CC	comprise only residues 383-706 or 383-713. The PHL catalyses the
CC	reaction: R-glyOH -> R-NH2. It can be used to produce peptides with
CC	amidated C-terminal, e.g. calcitonin, growth hormone, LH-RH.
XX	
SQ	Sequence 935 AA;
Query Match	4.9%; Score 100; DB 13; Length 935;
Best Local Similarity	19.4%; Pred. No. 0.47;
Matches	76; Conservative 57; Mismatches 148; Indels 110; Gaps 18;
Db	
33	CGPSTITNTFNTRNAEGHHVYVKGLYDDEGCGNDSGGRGVAGISLPDSCNVAFRSLNP 92
	: : : : :       : : :       : :       : :
526	CQPLDVAVDPITTGNEFVDGY-----CN-SRLMQESP 556
QY	93 RGIEV-----TTTIVTSFHPLEVTKYDRAVRVOCFMEADKTVS 131
	:   :   :   :   :   :   :   :   :   :   :
Db	557 NGFMFMQMGESLTSSNVRRPGGRIFRHSLSLWVPDQQLCVADENG-RIOCFIALEGNFVK 615
QY	132 AQIEVSEI-TTAFOQTQIYPMRYCRYELIDGEGTGOPYFAITGQPVYHRMTCDSETVDTF 190
	:   :   :   :   :   :   :   :   :   :   :
Db	616 -QIKHOEGREVFAYSAAP-----GG-----VLAVNGRPYY-----GX 648
QY	191 CAVVHSCVDDGNGGTVEILNADGCALDKYL-LNNLETPDDLMAGEAHVYKKYADRSL 248
	:   :   :   :   :   :   :   :   :   :   :
Db	649 SAPVGGEMLNFSNGD-----LDLTFPARKNEMPHDIAADGOTGVVGDHANAVA 698
QY	249 FYQCQIS-----ITIKEPN--SECVRPCQCEPOGAGVKTGAARKPALAARLIKLKRS 301
	:   :   :   :   :   :   :   :   :   :   :
Db	699 VMKSPSPAHERSVAKKAIQEVEITEETFEIHHSKRKTNSVAKQOEKOQK--QKNS 756
QY	302 A-----EPENIT-DVRDINTLE---ISDDNALFPVDLRNRALLQHNGOPVTLAAVUNGSI 352
	:   :   :   :   :   :   :   :   :   :   :
Db	757 AGVSTOEKONVVOEINAGVPTQOEKONVVOESAGVSTOEKGVSAGVSTOEKOSV 816
QY	353 CMSPFGSMFGSLIALI-AAVITISKFR 382
	:   :   :   :   :   :   :   :   :   :   :
Db	817 QESSAGVSEVLITLLIPIAVLIAIAIFIR 847
RESULT 22	
AAB71814	
ID	AAB71814 standard; Protein; 227 AA.
XX	

```

XX alpha-amidating; pAX799; alpha amide; ds.
KW Synthetic.
OS
XX EP299790-A.
PN
XX 18-JAN-1989.
PD
XX 15-JUL-1988; 88EP-0306508.
PF
XX 17-JUL-1987; 87JP-0177184.
PR
XX 05-DEC-1987; 87JP-0306867.
XX
XX (SUNR ) SUNTORX LTD.
PA
XX Ohsuye K, Kitano K, Tanaka S, Matsuo H, Mizuno K;
PI
XX WPI: 1989-017279/03.
DR
XX N-PSDB: AAN94527.
DR
XX
XX MPI: 1989-017279/03.
XX
XX N-PSDB: AAN94527.
XX
XX Recombinant C terminal alpha amidating enzymes of Xenopus laevis -
XX and their precursors deoxyribonucleic acid encoding sequences.
XX
XX Disclosure; 107pp; English.
XX
XX The sequence encodes a derivative of the mature C-terminal alpha-
XX amidating enzyme from plasmid pXA799.
XX The plasmid was screened from an E.coli library using plasmid pXA457
XX to screen a larger library.
XX Although pXA799 is similar to pXA457 at the N-terminus, it has an area
XX of hydrophobic elements suggesting a membrane function.
XX See also AAN93060.
XX
XX Sequence 693 AA:
SQ

```

Query Match 4.7%; Score 94.5; DB 10; Length 693;  
 Best Local Similarity 21.6%; Pred. No. 1.1;  
 Matches 50; Conservative 25; Mismatches 66; Indels 91; Gaps 12;

```

OY 33 CGPTSTITNFNTRNAFECHVYKGLYDQEGCRNDEGGRQVAGISLPFDSNVARTSLNP 92
DB 491 CQPTDVAVDPITGNFVADG-----CN-SRIMQSP 521
OY 93 RGEIV-----TTTVISFHPLEVTKYDRAYRVQCFYMEADKTVS 131
DB 522 NMFIMQGEETSNLPRGQFRIPHSITMISDQGLCVADRENG-RIOCFHAKGFEVK 580
OY 132 AQIEVSEI-TTAFQOIYMPYCRREILDGPTGQPVQFATIGQVYHKMTCDSFTVTF 190
DB 581 -QIKHQEFGREYFAVSYP-----VYAVNGKPYG---DSTPVQGF 620
OY 191 CAVVHSCFVDGNGDPTVEILNADGCALDKYL--LNNLEYPTDLMAGQEAHY 240
DB 621 -----MLNFSNGD-----ILDFTIPARKNFEMPHDIAAGDDGTYY 655

```

RESULT 24  
 AAP94854  
 ID AAP94854 standard; protein; 875 AA.  
 AC AAP94854;  
 XX  
 XX 27-JUN-1990 (first entry)  
 DT  
 XX C-terminal prepro-C-terminal alpha-amidating enzyme of pXA799.  
 DE  
 XX alpha-amidating; pAX799; alpha amide; ds.  
 KW  
 XX Synthetic.  
 OS  
 XX EP299790-A.  
 PN

```

PD 18-JAN-1989.
XX
XX 15-JUL-1988; 88EP-0306508.
PF
XX 17-JUL-1987; 87JP-0177184.
PR
XX 05-DEC-1987; 87JP-0306867.
XX
XX (SUNR ) SUNTORX LTD.
PA
XX Ohsuye K, Kitano K, Tanaka S, Matsuo H, Mizuno K;
PI
XX WPI: 1989-017279/03.
DR
XX N-PSDB: AAN90791.
DR
XX
XX MPI: 1989-017279/03.
XX
XX N-PSDB: AAN90791.
XX
XX Recombinant C terminal alpha amidating enzymes of Xenopus laevis -
XX and their precursors deoxyribonucleic acid encoding sequences.
XX
XX Disclosure; 107pp; English.
XX
XX Plasmid pXA799 contains a sequence derived from Xenopus laevis.
XX The plasmid was screened from an E.coli library using plasmid pXA457
XX to screen a larger library.
XX Although the gene product is similar to that of pXA457 at the
XX N-terminus, it has an area of hydrophobic elements suggesting a membrane
XX function. See also AAN93060.
XX
XX Sequence 875 AA:
SQ

```

Query Match 4.7%; Score 94.5; DB 10; Length 875;  
 Best Local Similarity 21.6%; Pred. No. 1.6;  
 Matches 50; Conservative 25; Mismatches 66; Indels 91; Gaps 12;

```

OY 33 CGPTSTITNFNTRNAFECHVYKGLYDQEGCRNDEGGRQVAGISLPFDSNVARTSLNP 92
DB 530 CQPTDVAVDPITGNFVADG-----CN-SRIMQSP 560
OY 93 RGEIV-----TTTVISFHPLEVTKYDRAYRVQCFYMEADKTVS 131
DB 561 NMFIMQGEETSNLPRGQFRIPHSITMISDQGLCVADRENG-RIOCFHAKGFEVK 619
OY 132 AQIEVSEI-TTAFQOIYMPYCRREILDGPTGQPVQFATIGQVYHKMTCDSFTVTF 190
DB 620 -QIKHQEFGREYFAVSYP-----VYAVNGKPYG---DSTPVQGF 659
OY 191 CAVVHSCFVDGNGDPTVEILNADGCALDKYL--LNNLEYPTDLMAGQEAHY 240
DB 660 -----MLNFSNGD-----ILDFTIPARKNFEMPHDIAAGDDGTYY 694

```

RESULT 25  
 AAR73053  
 ID AAR73053 standard; protein; 875 AA.  
 AC AAR73053;  
 XX  
 XX 06-NOV-1995 (first entry)  
 DT  
 XX Peptidyl C-terminal alpha-amidating enzyme.  
 DE  
 XX Peptidyl C-terminal alpha-amidating enzyme; AE; cell culture;  
 KW Trichostatin; CHO.  
 XX  
 XX Not specified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..39  
 FT /label= Sig\_peptide  
 PN  
 XX EP649900-A.  
 XX 26-APR-1995.  
 PD  
 XX 07-SEP-1994; 94EP-0306587.  
 PF









XX PS Claim 6; Page 113-114; 122pp; English.

XX CC The present sequence is the fusion protein construct, pIL2Trim3. This

CC construct comprises of the leader peptide from human Interleukin 2 (hIL2)

CC and soluble CD39 (solCD39) protein region, having a pyrolysate activity.

CC Soluble CD39 is constructed by removing the N- and C-terminal

CC transmembrane domains. It retains the capacity to metabolize ADP and ADP

CC at relevant concentrations and the ability to block and reverse

CC ADP-induced platelet activation and recruitment, including platelet

CC aggregation. Soluble CD39 polypeptides are useful for inhibiting

CC angiogenesis. It is useful for the treatment of unstable angina, stroke,

CC myocardial infarction, coronary artery disease or injury, embolism,

CC atherosclerosis, peripheral vascular occlusion, preclampsia, platelet-

CC associated ischaemic disorders including lung, coronary and cerebral

CC ischaemia, thrombotic disorders including coronary, peripheral and

CC cerebral artery thrombosis, intracardiac and venous thrombosis,

CC coagulopathy, deep venous thrombosis (DVT), pulmonary embolism (PE) and

CC transient ischaemic attack. Soluble CD39 is also useful for preventing

CC thrombus formation or reformation, occlusion, reocclusion, stenosis or

CC restenosis of blood vessels or stroke.

XX SQ Sequence 474 AA:

Query Match 4.3%; Score 87.5; DB 21; Length 474;

Best Local Similarity 20.9%; Pred. No. 3.5;

Matches 55; Conservative 40; Mismatches 75; Indels 93; Gaps 13;

QY 113 VDRAYVOCFYME-ADKTYSAQIEVEITPAPOTQVPMPCRYEILDGPGPGQPVQFAI 171

DB 5 IDRMQLSLCALSLALVYNSSTKTKQLTSSYONKALPENYKGYILDGSSHTSLYI-- 62

QY 172 IQOPVHKWTCDSFVDFCAVH-----SCFVDDGN----- 203

DB 63 -----YKMPAEKEN-DT--GVVHQVEECRGVKGPGISKVKQKNEIGYILDGMEARREY 113

QY 204 -----GDT--VEILNADGSCALDKYLNLNLE-----YPTD-----LMAGQEAH 238

DB 114 PRSCHOETPVYLGATAGRLRLMESEELADRVLDVERSLSNYPDFOGARITGQEGEG 173

QY 239 VY-----KYADRSOLFYOQCOISITIKPEPNSCVPRQSEPGCFQAVTGGAAAP 288

DB 174 AVGWITINVLKGFSGKTRWF-----STVPEYETNNQ-----ETFGALDGGASTQV 219

QY 289 AAAQRLKRRSAEPENIIDVR 311

DB 220 FVPPQNTIE-----SPDNALQFR 238

RESULT 31

AAAY70901

ID AAY70901 standard; Protein; 473 AA.

XX AAAY70901;

XX 17-AUG-2000 (first entry)

DE Protein encoded by Trim 4 construct.

XX Soluble CD39; ADP-induced platelet activation; platelet aggregation;

KW unstable angina; myocardial infarction; stroke; coronary artery disease;

KW atherosclerosis; peripheral vascular occlusion; preclampsia; embolism;

KW platelet-associated ischaemic disorder; lung ischaemia; thrombolytic;

KW cerebral ischaemia; thrombotic disorder; coronary artery thrombosis;

KW cerebral artery thrombosis; intracardiac thrombosis; venous thrombosis;

KW peripheral artery thrombosis; coagulopathy; deep venous thrombosis; DVT;

KW PE; pulmonary embolism; transient ischaemic attack; thrombus formation;

KW occlusion; reocclusion; stenosis; restenosis; antiangiinal; cardiac;

KW cerebroprotective; antiatherosclerotic; vasotropic; anticoagulant;

KW coronary ischaemia; vascular occlusion; pIL2Trim4 variant.

OS Homo sapiens.

OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Cleavage-site 26..27

XX FT Protein 35..473

XX FT note="Soluble portion of CD39"

PN MO200023094-A2.

XX 27-APR-2000.

XX 13-OCT-1999; 99MO-US23641.

XX 16-OCT-1998; 98US-0104585.

XX 06-NOV-1998; 98US-0107466.

XX 13-AUG-1999; 98US-0149010.

XX (IMV ) IMMUNEX CORP.

XX (CORR ) CORNELL RES FOUND INC.

XX Maliszewski CR, Gayle RB, Marcus AJ;

XX WPI; 2000-339518/29.

XX Inhibiting platelet activation and recruitment, useful for treating a

XX mammal suffering from unstable angina, myocardial infarction, stroke,

XX coronary artery disease or injury, comprises administering soluble CD39

XX polypeptides -

XX Example 11; Page 114-116; 118pp; English.

XX The present sequence is the protein encoded by Trim4 construct. pIL2Trim4

XX variant was constructed by removing the human IL2 residues from solCD39

XX fusion construct. Fusion of 12 amino acids from the N-terminus of mature

XX human IL2 to the solCD39 coding region results in high levels of both

XX expression and activity in the supernatants of transfected cells. SolCD39

XX is used in the treatment of unstable angina, myocardial infarction,

XX stroke, coronary artery disease or injury, atherosclerosis, peripheral

XX vascular occlusion, preclampsia, embolism, platelet-associated ischaemic

XX disorder including lung ischaemia, coronary ischaemia and cerebral

XX ischaemia, a thrombotic disorder including coronary artery thrombosis,

XX cerebral artery thrombosis, intracardiac thrombosis, peripheral artery

XX thrombosis, venous thrombosis, thrombosis, coagulopathy, deep venous

XX thrombosis (DVT), pulmonary embolism (PE), transient ischaemic attack.

XX Soluble CD39 is also useful for preventing thrombus formation or

XX reformation, occlusion, reocclusion, stenosis or restenosis of blood

XX vessels or stroke.

XX SQ Sequence 473 AA:

Query Match 4.3%; Score 87; DB 21; Length 473;

Best Local Similarity 19.5%; Pred. No. 4;

Matches 51; Conservative 42; Mismatches 77; Indels 92; Gaps 12;

QY 113 VDRAYVOCFYMEADKTYSAQIEVEITPAPOTQVPMPCRYEILDGPGPGQPVQFAI 172

DB 5 IDRMQLSLCALSLALVYNSSTKTKQLTSSYONKALPENYKGYILDGSSHTSLYI-- 61

QY 173 GOPYVHKWTCDSFVDFCAVH-----SCFVDDGN----- 203

DB 62 -----YKMPAEKEN-DT--GVVHQVEECRGVKGPGISKVKQKNEIGYILDGMEARREY 113

QY 204 -----GDT--VEILNADGSCALDKYLNLNLE-----YPTD-----LMAGQEAH 239

DB 114 PRSCHOETPVYLGATAGRLRLMESEELADRVLDVERSLSNYPDFOGARITGQEGEG 173

QY 240 Y-----KYADRSOLFYOQCOISITIKPEPNSCVPRQSEPGCFQAVTGGAAKPA 289

DB 174 XGWITINVLKGFSGKTRWF-----STVPEYETNNQ-----ETFGALDGGASTQV 219

QY 290 AAAQRLKRRSAEPENIIDVR 311

DB 220 FVPPQNTIE-----SPDNALQFR 237

Seq	Sequence	473 AA;	4.38;	Score 87;	DB 21;	Length 473;
Query Match	Best Local Similarity	19.58;	Pred. No. 4;			
Matches 51;	Conservative 42;	Mismatches 77;	Indels 92;	Gaps 12;		
QY	113 VDRARVROCFEMADKTVASQAEVSEITATROTQIVPAPCRVETILDGSGPTGQPVQFAII 172					
DB	5 IDRMQLISCLALSTALVTNSSTKTKLTSSYONKRALPENVYGGIVLDGSSHTSLYT--- 61					
QY	173 GQPVYHKWTCDESEIVDFECAVH-----SCEVDGDN----- 203					
DB	62 -----YKMPAEKEN-DT--GVYHVEEBCRVAGPGSKRVQKVNENIGIYLTOMERAREVI 113					
QY	204 -----GDT--VEILNADGALDKYLLNNLE-----YPTD-----LMAGQEAHV 239					
DB	114 PRSOHETPVYILGATAGMRLLRMESEELADRLDVERSLSNYPDPFGARHITGOEBGA 173					
QY	240 Y-----KYADRSGLFYQCQISITIKRPNSECVAPQCSPEQGFPAVKTGGAAAPRA 289					
DB	174 YGWTITVYLLGKFSOKTRFMF-----SIVPEYTNNO-----ETFGALDGGASTGYT 219					
QY	290 AAOLRLKRSAPENIIDVR 311					
DB	220 FVPONQITE----SPDNALQFR 237					
RESULT 33						
AAAR741171						
ID	AAAR741171 standard; Protein: 3038 AA.					
AC	AAAR741171;					
DT	18-JAN-1996 (first entry)					
DE	Aspergillus terreus triol polyketide synthase.					
KW	Triol polyketide synthase; TPKS; HMG-CoA reductase inhibitor;					
KM	hypercholesterolaemia; LDL-cholesterol.					
OS	Aspergillus terreus.					
XX						
XX	Key	Location/Qualifiers				
FT	Region	181				
FT	Region	/label= keto-acyl synthase motif				
FT	Region	654..658				
FT	Region	/label= acetyl/malonyl transferase motif				
FT	Region	985..994				
FT	Region	/label= dehydratase motif				
FT	Region	1446..1450				
FT	Region	/label= methyl transferase motif				
FT	Region	1932..1937				
FT	Region	/label= enoyl reductase motif				
FT	Region	2164..2169				
FT	Region	/label= keto reductase motif				
FT	Region	2498				
FT	Region	/label= acyl carrier protein motif				
FT	Misc-difference	282..288				
FT	Misc-difference	/label= misc feature				
FT	Misc-difference	1450..1460				
FT	Misc-difference	/label= misc feature				
FT	Misc-difference	1603..1612				
FT	Misc-difference	/label= misc feature				
FT	Misc-difference	2521..2535				
FT	Misc-difference	/label= misc feature				
XX						
XX	WO9512661-A.					
XX						
XX	11-MAY-1995.					
XX						
XX	28-OCT-1994;	94WO-US12423.				
XX						
XX	02-NOV-1993;	93US-0148132.				
XX						

```

XX (MERI ) MERCK & CO INC.
PA
XX
XX Conder MJ, Davis CR, Hendrickson LE, Macada PC, Rambosek J;
PI Reeves CD, VincI VA;
XX
XX WPI; 1995-193816/25.
DR N-PSDB; AA092323.
XX
PT Novel DNA encoding triol poly-ketide synthase - used to isolate and
PT identify homologues of triol poly-ketide synthase, and in the treatment
PT of hyper-cholesterolaemia
XX
XX Claim 12; Figure 2; 107pp; English.
PS
XX
XX The full-length TPKS-encoding DNA in plasmid pLOA was
XX designated pPKS100. Splicing of the introns from the DNA
XX sequence and translation of the 914 nt ORF results in a
XX protein of 3038 AAs (AA074171) with a mol. wt. of 269,090
XX daltons. Inspection of the TPKS AA sequence for active
XX site residues and motifs known to be associated with
XX polyketide synthases and fatty acid synthase (FAS) activities
XX resulted in the identification of candidates for expected
XX sites (see FT). Except for the presence of a methyl transferase,
XX not present in FAS, the succession of activities on the
XX TPKS protein is the same as that observed for the rat FAS
XX protein.
XX
XX
XX Sequence 3038 AA;
SQ
Query Match 4.3%; Score 87; DB 16; Length 3038;
Best Local Similarity 19.7%; Pred. No. 77;
Matches 93; Conservative 55; Mismatches 138; Indels 186; Gaps 24;
QY 6 IACCTTILALSYSPVDNGVEPEIEIC--GPTSI-----TINFT 44
DB 1175 VAOQYIGA--YSSP-----GDRRLCLYVPHVDRIPLSLCLATAESGCEKAFNT 1226
QY 45 RNFECHVYVYK---LYDDEGCHNDGSGHVGISL--PDSCHVATRSINPGIR--V7 98
DB 1227 INTYDGDYLSGDIYVFDAE---QTTLOVENTITKPPSPASVDHAFHAFMSWGPIIT 1282
QY 99 TTVVVISFHLFVTKVDR-----AYRQCFYMEADKTVASQIEVSEITTA---FQTOIVPM 150
DB 1283 PDSILNDPEXWATAQDKKAIPITERIVFYI---RSFLQDLTEEGQAFLHQKIEWL 1339
QY 151 PVCRYEILDGGPTGPGPYQAIIGQPYVHKWTCDSFVDFFCAVHSCFVDD----- 201
DB 1340 E---OVLASAKEGR-----HLMYDPGWENDTEAOIEHLCTANSYHPHRLVQ 1383
QY 202 -----GNGDTVEILNDGALDKYLLNN-----LEPYDIDMAGQEAHYKXAD 244
DB 1384 RVGQHLLPVRNSNGNFEDLLDHG--LITEFYNTLISFGALHYARLVA--QIAHRYQSM 1441
QY 245 RSOLFVQCOISITIKRPNSECVRPQCSFPGCAVKTGAARKPAAQAOL----- 294
DB 1442 ILEI-----GAG---TGATYVVALTPQLGNSITYYDI 1472
QY 295 -----RLKKRSAPENIIDVRDINTLEISDDNOLP 327
DB 1473 STGFEEQAREQAPFEDRKYVEPLDIRSPAQGEF--HAYDLIIASNVLAHTPD---LE 1528
QY 328 VDLRH-RALLQHNQGPVIL-----AAVONGICMSPF 357
DB 1529 KTMHAHRSILKPGQGVILEITHKHETRLGLTFGLFADMMAGVDDGRCTEPF 1580

```

```

RESULT 34
AAB94352
ID AAB94352 standard; Protein; 636 AA.
XX
AC AAB94352;
XX

```

```

DT 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:14870.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
PN
XX
XX 07-FEB-2001.
PD
XX
XX
XX 28-JUL-2000; 2000EP-0116126.
PF
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 14870; 2537bp + CD ROW; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX
XX Sequence 636 AA;
SQ
Query Match 4.3%; Score 86.5; DB 22; Length 636;
Best Local Similarity 21.9%; Pred. No. 7.3;
Matches 84; Conservative 54; Mismatches 153; Indels 93; Gaps 23;
QY 6 IACCTTILALSYSPVD--NGVEGEPEIEICGPTSTITINNTNANFEGHYVVGGLDQBCR 64
DB 259 VGFLSSLLPQSKKSPSRKSPAQGPPO---PQSSA---KRESFGQ--GTGKDKPTSGAK 309
QY 65 NDEGROYAGISLPFDSCHVATRSINRG-----IEVTTVVVISFHLFVTKVDRAYRV 119
DB 310 --DGKSLSLGATGESGSHQRQRRLDQHKERKEIPSTTT-----S 349
QY 120 QCFYMEADKTVASQIEVSEITTAFFQTOIVPMPCRYEILDGGPTG---QPYQ--FAITG 175

```

```
Db 350 QC-----AEKKPEASGPEAPECPPELHTEPVE-PLTRAS--SAGPRGGGVREPPFIVLGOE 402
Oy 176 VYHKWTCDSFVDPFCFAVHSCFVDGNGDVEILNADG-----CALDKY 220
Db 403 EYG-----EHHSSIMH-CRY-DCSGRRVASLDVDGYIKWSEFNPIMOTKASSISKS 451
Oy 221 LNNLEYPD---LMAGOEAHVYKYADRSOLFQCOISITIKPEPNECVAPQCSEPOGF 276
Db 452 PLTLEWATKRDRLLLLGSGVGYRLYDTEAKKMLCEININDMP--RIISLACS-PNGA 508
Oy 277 GAVKGTGAAPAAAOQLRL-----LKKRSAPENIT--DVRTDINTLEISDONQALPV 328
Db 509 SFV---CSAAPSLTSGVDESAPDIGSKMNOVPGRLLMDTKTKMQQLQFSLDPEPIAI 565
Oy 329 DLRRHALLQHNGOPVITLAAYONGI 352
Db 566 NC---TAFNHNGNLVTGAADGVI 586

RESULT 35
AAM40070
ID AAM40070 standard; Protein; 747. AA.
AC AAM40070;
XX
XX
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 3215.
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PA Tang YT, Iau C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
DR N-SDb; AAI59226.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 5; SEQ ID NO 3215; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
```

```
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 747 AA;
SQ
Query Match 4.3%; Score 86.5; DB 22; Length 747;
Best Local Similarity 21.9%; Pred. No.9.4; Indels 93; Gaps 23;
Matches 84; Conservative 54; Mismatches 153;
Oy 6 IAFCTTLTALSYSIPVD-NGVEGEPEIECPSTITINFNTNAPFGHYVYGLYDOEGCR 64
Db 259 VGFLSSLLPQSKSPSRISPAQGPQ---PQSSA---KKESEFGQ-CTGGRDPTSGAK 309
Oy 65 NDEGGROYAGISLPFDSQNVARTSLNDRG-----IFVTTTVVISFHPLEVTKYDRAVRV 119
Db 310 --DGKSLSLGLATGESGWSQHRQRLQDHGKERKELEFSTTT-----S 349
Oy 120 QCFYMEADKTVSAQIEVSEITTAFOFOIVPMPVCRVETLDDGPRG---QPVQ-PAITGQP 175
Db 350 QC-----AEKKPEASGPEAPECPPELHTEPVE-PLTRAS--SAGPRGGGVREPPFIVLGOE 402
Oy 176 VYHKWTCDSFVDPFCFAVHSCFVDGNGDVEILNADG-----CALDKY 220
Db 403 EYG-----EHHSSIMH-CRY-DCSGRRVASLDVDGYIKWSEFNPIMOTKASSISKS 451
Oy 221 LNNLEYPD---LMAGOEAHVYKYADRSOLFQCOISITIKPEPNECVAPQCSEPOGF 276
Db 452 PLTLEWATKRDRLLLLGSGVGYRLYDTEAKKMLCEININDMP--RIISLACS-PNGA 508
Oy 277 GAVKGTGAAPAAAOQLRL-----LKKRSAPENIT--DVRTDINTLEISDONQALPV 328
Db 509 SFV---CSAAPSLTSGVDESAPDIGSKMNOVPGRLLMDTKTKMQQLQFSLDPEPIAI 565
Oy 329 DLRRHALLQHNGOPVITLAAYONGI 352
Db 566 NC---TAFNHNGNLVTGAADGVI 586

RESULT 36
AAB93124
ID AAB93124 standard; Protein; 747. AA.
AC AAB93124;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12006.
XX
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
```



XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 12006; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-qt primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
XX Sequence 747 AA:  
SQ  
Query Match 4.3%; Score 86.5; DB 22; Length 747;  
Best Local Similarity 21.9%; Pred. No. 9.4;  
Matches 84; Conservative 54; Mismatches 153; Indels 93; Gaps 23;  
QY 6 IACTTTLALSYSLPVD-NQVEGEPELEGPTSTITNFNRMNAPFEGHYVYKGLXDQGR 64  
DB 259 VGLSSLPLPSKSPSPRLSPAQGPQ---PQSSA---KKESEFGQ-GTKGRDPTSGAK 309  
QY 65 NDEGQVAGISLPFDSNCNARTRSINPRG---IFVTTTVISFHPLEFVKVDRAYRV 119  
DB 310 --DGKSLSLGATGESGWSGSHRRLQDHKKEKELFTTT-----S 349  
QY 120 QCFYMEADKTVSAQIEVSEITTAFOQIYPMPCRYEILIDGPTG---QPVQ-FAITGQP 175  
DB 350 QC---AEKPEASGPEAEPCPELHTEPVE-PLTRAS--SAGPEGGVREQGFYVLQGE 402  
QY 176 VYHKWTCDSFVDTFCVAVHSCFVDGNGDVTVELNADG-----CALDKY 220  
DB 403 EYV-----EHHSSIMH-CRV-DCSGRRVASLDVDGVIKWSFNPITQTKASSISKS 451  
QY 221 LNNLELYPTD---LMAQGEAHYKYXADRSQLEFYOCQISITIKPEPNSCEVRPCSPQGE 276  
DB 452 PLILSEMAKRRRLLLIGSGVGVRLYDTEAKNLCENINDMP--RLISLACS-PNGA 508  
QY 277 GAVKTTGGAAPAAQAOLRL-----LKKRSAPENII--DVRTDINTLEISDDNOLAVY 328  
DB 509 SFV---CSAAPFELTSQVFSAPDIDSGKGNQVGRLLMLDPTKMQOQLQFSLDPPRIAI 565  
QY 329 DLRRHALLQHNQPVITLAVQNKI 352  
DB 566 NC---TAFNHNGMLVTGADGVI 586

RESULT 37  
ABP38294

ID ABP38294 standard; Protein; 438 AA.  
XX  
XX AC ABP38294;  
XX  
XX DT 24-JUL-2002 (first entry)  
XX  
XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3139.  
XX  
XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
KW antibacterial; gene therapy.  
XX  
XX OS Staphylococcus epidermidis.  
XX  
XX PN US6380370-B1.  
XX  
XX PD 30-APR-2002.  
XX  
XX PF 13-AUG-1998; 98US-0134001.  
XX  
XX PR 14-AUG-1997; 97US-055779P.  
XX  
XX PR 08-NOV-1997; 97US-064964P.  
XX  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX PI Doucette-Stamm LA, Bush D;  
XX  
XX DR WPI; 2002-381255/41.  
XX  
XX DR N-PADB; ABN90839.  
XX  
XX PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
XX polypeptide, useful for diagnosing and treating bacterial infections -  
XX  
XX PS Disclosure; SEQ ID 3139; 267pp; English.  
XX  
XX CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
XX frame (ORF) nucleic acid sequences which encode the amino acid sequences  
XX given in ABP35124 to ABP37960. The S. epidermidis sequences have  
XX antibacterial activity and can be used in gene therapy. The sequences  
XX can also be used in the diagnosis and treatment of bacterial infections,  
XX particularly S. epidermidis infections. The sequences can be used to  
XX screen for compounds able to interfere with the S. epidermidis life  
XX cycle or inhibit S. epidermidis infection.  
XX  
XX CC N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from the  
XX USPTO web site.  
SQ  
Sequence 438 AA:  
SQ  
Query Match 4.2%; Score 85.5; DB 23; Length 438;  
Best Local Similarity 19.2%; Pred. No. 5.1;  
Matches 62; Conservative 49; Mismatches 127; Indels 85; Gaps 13;  
QY 32 ECGPTSTITNFNRMNAPFEG-HYVYKGLXDQGRNDEGGRQVAGSLPFDSCNARTRS 89  
DB 154 ENGFTKIDPILITASAPETSELEFHTKYDEDAFLSQS-----QLYMEAAAANGH- 205  
QY 90 LMRGIFVTTVVVISHPLEFVKVDRAYR--YOCFYMEADKTVSAQIEVSEITTAFOQI 147  
DB 206 -----VFSGPFRAEKSTRRHLLIEFWMIEPMAFTNHAESLEQEOYVSHI 253  
QY 148 VPMV--CRYEILIDGPTGQPVQFALIGQVYHKWTCDSFVDTFCVAVHSCFVDGNGD 205  
DB 254 VOSVLNHCQLEIK-----ALDRDT--TKLEKVAFTFPRISSYD 289  
QY 206 TVEILNADQCALDKYLLNNLEFPTDLMAQGEAHYKYXADRSOLF--YOCQISITIKEPS 263  
DB 290 ATEFLKKEG-----FDIEWGEDGFGAPHEIALNHHYDLVPFTNNYPTKIKPFYMQPMP 342  
QY 264 E-----CVRPCSEPOGFAVKTGGAAPAAQAOLRLKKRSAPENIIDVRTDINTL 317  
DB 343 DNEEDTVLCA--DLIAPBGGEIIGSERIN-----DLELLEOR-----INEH 382  
QY 318 EISDDNQLPVDLRRHALLQHNQ 340

Db 383 ELDESSSYLDLRKRGVPHSG 405

## RESULT 38

AAAM13009  
ID AAAM13009 standard; protein: 560 AA.

XX AAAM13009;

XX 21-NOV-1997 (first entry)

XX Segment of desmosomal cadherin, desmoglein Dsg2.

XX Desmosomal cadherin; desmoglein, Dsg2; cell; surface; epithelial;  
KW Carcinoma; desmosome; antibody; epitope; diagnosis; detection;  
KW micrometastasis; separation; enrichment; targeted delivery;  
KW metastatic.

XX Homo sapiens.

XX DE19531033-A1.

XX 27-FEB-1997.

XX 23-AUG-1995; 95DE-1031033.

XX 23-AUG-1995; 95DE-1031033.

XX (PROG-) PROGEN BIOTECHNIK GMBH.

XX Franke WW, Schaefer S;

XX WPI; 1997-146518/14.

XX Antibody reactive with part of desmosomal cadherin - exposed on  
PT surface of epithelial or carcinoma cells, not bound to desmosomes,  
PT useful for diagnosis and treatment of carcinoma micrometastases

PS Claim 7; Page 5; 8pp; German.

XX The present sequence is a segment of the desmosomal cadherin (DC),  
CC desmoglein Dsg2, which is exposed on the surface of epithelial or  
CC carcinoma cells and not bound to desmosomes. An antibody (Ab)  
CC directed against epitopes of the present sequence can be used to  
CC diagnose, i.e. to detect carcinoma cells, especially  
CC micrometastases, not bound to desmosomes, to separate, enrich or  
CC detect living or fixed carcinoma cells by cell sorting methods and  
CC as a therapeutic to deliver agents, e.g. other Ab or toxins, to  
CC target cells. The Ab provides rapid and reliable detection of  
CC metastatic carcinoma, and detects parts of DC that are not  
CC accessible in desmosome bound cells, as in normal tissue or  
CC carcinomas.

XX Sequence 560 AA;

XX Query Match 4.2%; Score 85.5; DB 18; Length 560;

XX Best Local Similarity 25.6%; Pred. No. 7.6;

XX Matches 5; Conservative 27; Mismatches 81; Indels 43; Gaps 12;

XX 15 LSYSPVNGVEGEPEI-ECGPTSTINENTRNAFEG-----HYVYGLYDQEGCR 64

XX 299 LDFEVIYANKRAKAFHDSIRSKKPPPIKVKVKNKBGIIHKSSVISIYSESMD---R 354

XX 65 NDEGGROYAGISLPDSCN---VARTSLNPRGIF-----VTTVVISFHPLEVTKY-- 113

XX 355 SSKG--GIIGNFOAFDEDTGLPAHARYKLEDRDNMISVDSVTSEIKLAKLPDESRVYQ 412

XX 114 DRAYRQCFTWMEAD---KTVSAQ---IEVSETTAFOQOIVM-PVCR-----YELTD 159

XX 413 NGYTVKIVASIEDYPRKTTIGVLTINVEDINDNCPTLIEPVQTCDAEYVNTAEDLD 472

XX 160 GGPTGPQVQFAIIGOP--VYHKW 180

Db 473 GHPNSGPFSEFSVIDKPPGMARKW 495

## RESULT 39

AAAM39436  
ID AAAM39436 standard; Protein: 940 AA.

XX AAAM39436;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2581.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW Peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW Leukemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-0534263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX WPI: 2001-442253/47.

XX N-PSDB: AA158592.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX Example 4; SEQ ID NO 2581; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAAM38642-AAAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activation/inhibition activity, chemotactic/chemokinetic activity, hemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX Sequence 940 AA;

XX Query Match

XX Best Local Similarity 4.2%; Score 85.5; DB 22; Length 940;

XX 25.6%; Pred. No. 17;



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 08:51:14 ; Search time 18 Seconds  
(without alignments)  
632.593 Million cell updates/sec

Title: US-10-054-562a-4

Perfect score: 2026

Sequence: 1 MMRLIAFCYTLIALSYSP.....ALIAAVITITSPKRNQKA 387

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents-AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/5A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/5B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfilist1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2026	100.0	387	4	US-09-323-427-4
2	2026	100.0	387	4	US-09-812-642-4
3	1086.5	53.6	245	4	US-09-323-427-17
4	1086.5	53.6	245	4	US-09-812-642-17
5	776.5	38.3	271	4	US-09-323-427-9
6	776.5	38.3	271	4	US-09-812-642-9
7	100	4.9	935	2	US-07-707-367-2
8	96.5	4.8	989	2	US-08-070-301-16
9	87	4.3	3038	1	US-08-450-332-2
10	87	4.3	3038	2	US-08-637-640-2
11	87	4.3	3038	4	US-09-004-406C-2
12	85.5	4.2	438	4	US-09-134-001C-3139
13	83	4.1	1222	2	US-08-682-517-15
14	83	4.1	1252	2	US-08-682-517-15
15	81.5	4.0	1911	1	US-08-348-0068-5
16	81.5	4.0	1911	2	US-08-800-825A-5
17	81.5	4.0	1911	4	US-09-158-657-5
18	81.5	4.0	1911	5	PCT-US94-10166-5
19	80.5	4.0	430	1	US-08-785-076-3
20	80.5	4.0	430	4	US-09-018-824-3
21	80.5	4.0	432	1	US-08-785-076-2
22	80.5	4.0	432	4	US-09-018-824-2
23	78.5	3.9	434	2	US-08-484-575A-20
24	78.5	3.9	434	3	US-08-477-459-20
25	78.5	3.9	434	3	US-08-479-869-20
26	78.5	3.9	434	4	US-08-486-414-46
27	78.5	3.9	434	5	PCT-US94-01826A-20

28	78.5	3.9	434	5	PCT-US94-02252A-20	Sequence 20, Appl
29	78.5	3.9	434	5	PCT-US96-03916-11	Sequence 11, Appl
30	78.5	3.9	434	5	PCT-US96-03916-67	Sequence 67, Appl
31	78.5	3.9	570	2	US-08-484-593B-16	Sequence 16, Appl
32	78.5	3.9	570	2	US-08-484-158B-16	Sequence 16, Appl
33	78.5	3.9	570	2	US-08-484-596A-16	Sequence 16, Appl
34	78.5	3.9	570	2	US-08-480-150A-16	Sequence 16, Appl
35	78.5	3.9	570	3	US-08-458-731-16	Sequence 16, Appl
36	78.5	3.9	570	3	US-08-149-223A-16	Sequence 16, Appl
37	78	3.8	549	1	US-08-325-071-61	Sequence 61, Appl
38	78	3.8	549	4	US-08-461-004A-61	Sequence 61, Appl
39	78	3.8	650	1	US-08-325-071-63	Sequence 63, Appl
40	78	3.8	650	4	US-08-461-004A-63	Sequence 63, Appl
41	77.5	3.8	650	4	US-09-336-643A-10	Sequence 10, Appl
42	77.5	3.8	650	4	US-08-325-071-67	Sequence 67, Appl
43	77.5	3.8	650	4	US-08-461-004A-67	Sequence 67, Appl
44	77	3.8	4536	4	US-09-180-422B-27	Sequence 27, Appl
45	76.5	3.8	1184	2	US-08-918-914-1	Sequence 1, Appl

## ALIGNMENTS

## RESULT 1

US-09-323-427-4  
; Sequence 4, Application US/09323427  
; Patent No. 6248329  
; GENERAL INFORMATION:  
; APPLICANT: Chandrasekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticulin Proteins, Nucleic Acid  
; TITLE OF INVENTION: Molecules, and Uses Thereof  
; FILE REFERENCE: HR-8  
; CURRENT APPLICATION NUMBER: US/09/323.427  
; CURRENT FILING DATE: 1999-06-01  
; EARLIER APPLICATION NUMBER: 60/087,435  
; EARLIER FILING DATE: 1998-06-01  
; NUMBER OF SEQ. ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 387  
; TYPE: PRT  
; ORGANISM: Dirofilaria immitis  
US-09-323-427-4

## Query Match

Best Local Similarity 100.0%; Score 2026; DB 4; Length 387;

Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MMRLIAFCYTLIALSYSPVNDNGVEGEPEIECGPTSTITNFNRNMFEGHVYKGLYDQ	60
DB	1	MMRLIAFCYTLIALSYSPVNDNGVEGEPEIECGPTSTITNFNRNMFEGHVYKGLYDQ	60
QY	61	EGCNDGGRVAGISLPFDSNVARTRSLNPRGIEFTTVVISFHPLEVTKVDRAVRQ	120
DB	61	EGCNDGGRVAGISLPFDSNVARTRSLNPRGIEFTTVVISFHPLEVTKVDRAVRQ	120
QY	121	CFYEMADKTSYSAQLEVESEITTAFTQTVPMPCVRCYETIIDDGPTGQPVQFALITGQPVYHKW	180
DB	121	CFYEMADKTSYSAQLEVESEITTAFTQTVPMPCVRCYETIIDDGPTGQPVQFALITGQPVYHKW	180
QY	181	TCDSYVDFTCVAVVHSCFVDDGNGDVEIILNADSCALDKYLILNNLEYPTDLMAQGEAHVY	240
DB	181	TCDSYVDFTCVAVVHSCFVDDGNGDVEIILNADSCALDKYLILNNLEYPTDLMAQGEAHVY	240
QY	241	KYADRSLFYQOCQSTITKEPNSCVRPQCSFPGFCAVVTGGAAPAAAOURLKKR	300
DB	241	KYADRSLFYQOCQSTITKEPNSCVRPQCSFPGFCAVVTGGAAPAAAOURLKKR	300
QY	301	SAEPENIIVRTDINTLEISDDNQALPVDLRHRALLQHNQOPVILAVONGICSPGFS	360
DB	301	SAEPENIIVRTDINTLEISDDNQALPVDLRHRALLQHNQOPVILAVONGICSPGFS	360

Qy 361 MEMGSIALLAAVITTSKFRPNOKA 387  
Db 361 MEMGSIALLAAVITTSKFRPNOKA 387

## RESULT 2

US-09-812-642-4  
; Sequence 4, Application US/09812642  
; Patent No. 6368600  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HM-8  
; CURRENT APPLICATION NUMBER: US/09/812,642  
; PRIOR FILING DATE: 1999-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 387  
; TYPE: PRT  
; ORGANISM: Dirofilaria immitis  
US-09-812-642-4

Query Match 100.0%; Score 2026; DB 4; Length 387;  
Best Local Similarity 100.0%; Pred. No. 4,3e-231;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMIRLAFCTTLIALSY-SPVNDNGVEGEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 60  
Db 1 MMIRLAFCTTLIALSY-SPVNDNGVEGEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 60  
Qy 61 EECRDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 120  
Db 61 EECRDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 120  
Qy 121 CFYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 180  
Db 121 CFYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 180  
Qy 181 TCDSETVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 240  
Db 181 TCDSETVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 240  
Qy 241 KYADRSOLFYOQOISITTIKEPNSCEVRPOCSEPOGFGAAYKTGAAPAAQOLRLKKR 300  
Db 241 KYADRSOLFYOQOISITTIKEPNSCEVRPOCSEPOGFGAAYKTGAAPAAQOLRLKKR 300  
Qy 301 SAEPENIIDVTDINTLEISDNOALPYDLRHRALLQHNQGPVILAAVONGICMSPFGFS 360  
Db 301 SAEPENIIDVTDINTLEISDNOALPYDLRHRALLQHNQGPVILAAVONGICMSPFGFS 360  
Qy 361 MEMGSIALLAAVITTSKFRPNOKA 387  
Db 361 MEMGSIALLAAVITTSKFRPNOKA 387

## RESULT 3

US-09-323-427-17  
; Sequence 17, Application US/09323427  
; Patent No. 6248329  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HM-8  
; CURRENT APPLICATION NUMBER: US/09/323,427  
; PRIOR FILING DATE: 1999-06-01  
; EARLIER APPLICATION NUMBER: 60/087,435

; EARLIER FILING DATE: 1998-06-01  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Brugia malayi  
US-09-323-427-17

Query Match 53.6%; Score 1086.5; DB 4; Length 245;  
Best Local Similarity 80.6%; Pred. No. 3,8e-120;  
Matches 195; Conservative 24; Mismatches 22; Indels 1; Gaps 1;

Qy 3 IRLIAFCTTLIALSY-SIPVNDNGVEGEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 61  
Db 4 MOICFSLSYMIASINAIPIIDNGVESEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 63  
Qy 62 GCRNDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 121  
Db 62 GCRNDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 121  
Qy 122 FYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 181  
Db 122 FYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 181  
Qy 182 CDSEIVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 241  
Db 182 CDSEIVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 241  
Qy 242 YA 243  
Db 244 YA 245

## RESULT 4

US-09-812-642-17  
; Sequence 17, Application US/09812642  
; Patent No. 6368600  
; GENERAL INFORMATION:  
; APPLICANT: Chandrashekar, Ramaswamy  
; APPLICANT: Morales, Tony H.  
; TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid  
; FILE REFERENCE: HM-8  
; CURRENT APPLICATION NUMBER: US/09/812,642  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 09/323,427  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Brugia malayi  
US-09-812-642-17

Query Match 53.6%; Score 1086.5; DB 4; Length 245;  
Best Local Similarity 80.6%; Pred. No. 3,8e-120;  
Matches 195; Conservative 24; Mismatches 22; Indels 1; Gaps 1;

Qy 3 IRLIAFCTTLIALSY-SIPVNDNGVEGEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 61  
Db 4 MOICFSLSYMIASINAIPIIDNGVESEPEIECGPTSTINFTNRAFEHGYVVKGLYDQ 63  
Qy 62 GCRNDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 121  
Db 62 GCRNDEGGRQVAGISLPDSCNVAARTSLNPRGIFVTTTVISFHPLEVTVDRAVRYO 121  
Qy 122 FYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 181  
Db 122 FYMEADKTVSAQIEVSEITTAFOQIYVMPVCRYEILIDGPTGQPVQFAIIGOPVYHMK 181  
Qy 182 CDSEIVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 241  
Db 182 CDSEIVDFCAVHSCFVDGNGDTVELNADGALDKYLLNNLEYPTDLMAGOEAHVK 241

```

Db      184 CDSEVTFFCALVHSCFVDDGNGDSINLINEGCAIDRLINLEIYFTDLMAGEAHYK 243
      |||||:|||||:::|||||:|||||
Qy      242 YA 243
      ||
Db      244 YA 245

```

```

RESULT 5
US-09-323-427-9
: Sequence 9, Application US/09323427
: Patent No. 6248329
: GENERAL INFORMATION:
: APPLICANT: Chandrashekar, Ramaswamy
: APPLICANT: Morales, Tony H.
: TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid
: TITLE OF INVENTION: Molecules, and Uses Thereof
: FILE REFERENCE: HW-8
: CURRENT APPLICATION NUMBER: US/09/323,427
: CURRENT FILING DATE: 1999-06-01
: EARLIER APPLICATION NUMBER: 60/087,435
: EARLIER FILING DATE: 1998-06-01
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 271
: TYPE: PRN
: ORGANISM: Dirofilaria immitis
: US-09-323-427-9

```

[illegible]

```

RESULT 6
US-09-812-642-9
: Sequence 9, Application US/09812642
: Patent No. 6368600
: GENERAL INFORMATION:
: APPLICANT: Chandrashekar, Ramaswamy
: APPLICANT: Morales, Tony H.
: TITLE OF INVENTION: Parasitic Helminth Cuticlin Proteins, Nucleic Acid
: FILE REFERENCE: MW-8
: CURRENT APPLICATION NUMBER: US/09/812,642
: CURRENT FILING DATE: 2001-03-20
: PRIOR APPLICATION NUMBER: 09/323,427
: PRIOR FILING DATE: 1999-06-01
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 271
: TYPE: PRN
: ORGANISM: Dirofilaria immitis
: US-09-812-642-9

```

[illegible]

```

1      RESULT 7
2      US-07-707-367-2
3      : Sequence 2, Application US/07707367
4      : Patent No. 5196316
5      : GENERAL INFORMATION:
6      :   APPLICANT: Iwasaki, Yasuno
7      :   APPLICANT: Shimoi, Hiroko
8      :   APPLICANT: Suzuki, Kenji
9      :   APPLICANT: Ghisalba, Oreste
10     :   APPLICANT: Nishikawa, Yoshiki
11     :   APPLICANT: Kawahara, Takashi
12     :   APPLICANT: Kangawa, Kenji
13     :   TITLE OF INVENTION: No. 5196316el Enzyme and DNA coding Therefor
14     :   NUMBER OF SEQUENCES: 2 _
15     :   CORRESPONDENCE ADDRESS:
16     :   ADDRESSEE: CIBA-GEIGY Corporation
17     :   STREET: 7 Skyline Drive
18     :   CITY: Hawthorne
19     :   STATE: New York
20     :   COUNTRY: USA
21     :   ZIP: 10532
22     :   COMPUTER READABLE FORM:
23     :   MEDIUM TYPE: Floppy disk
24     :   COMPUTER: IBM PC compatible
25     :   OPERATING SYSTEM: PC-DOS/MS-DOS
26     :   SOFTWARE: PatentIn Release #1.0, Version #1.25
27     :   CURRENT APPLICATION DATA:
28     :   APPLICATION NUMBER: US/07707,367
29     :   FILING DATE: 19910530
30     :   CLASSIFICATION: 530
31     :   PRIOR APPLICATION DATA:
32     :   APPLICATION NUMBER: JP 141678/90
33     :   FILING DATE: 01-JUN-1990
34     :   PRIOR APPLICATION DATA:
35     :   APPLICATION NUMBER: JP 210535/90
36     :   FILING DATE: 10-AUG-1990
37     :   PRIOR APPLICATION DATA:
38     :   APPLICATION NUMBER: JP 329911/90
39     :   FILING DATE: 30-NOV-1990
40     :   ATTORNEY/AGENT INFORMATION:
41     :   NAME: Williamizar, Johan
42     :   REGISTRATION NUMBER: 30,598
43     :   REFERENCE/DOCKET NUMBER: 4-18110/A/CGJ 44
44     :   TELECOMMUNICATION INFORMATION:
45     :   TELEPHONE: (914)785-7120
46     :   TELEFAX: (919)541-8689
47     :   INFORMATION FOR SEQ ID NO: 2:
48     :   SEQUENCE CHARACTERISTICS:
49     :   LENGTH: 935 amino acids
50     :   TYPE: AMINO ACID
51     :   TOPOLOGY: linear
52     :   MOLECULE TYPE: protein

```







QY 328 VDLRH-RALLQHNQPVIL-----AAVONGICMSPF 357  
 DB 1529 KTMAMARSILKPGGQVILITHTKHTRLGIFGLFADWAGVDDGRCTEPF 1580

## RESULT 11

US-09-004-406C-2  
 ; Sequence 2, Application US/09004406C  
 ; Patent No. 6174706  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vincel, Victor A.  
 ; APPLICANT: Conder, Michael J.  
 ; APPLICANT: McAda, Phyllis C.  
 ; APPLICANT: Reeves, Christopher D.  
 ; APPLICANT: Rambosek, John  
 ; APPLICANT: Davis, Charles Ray  
 ; APPLICANT: Hendrickson, Lee E.  
 ; TITLE OF INVENTION: DNA ENCODING TRIOL POLYKETIDE SYNTHASE  
 ; FILE REFERENCE: 19076FDA  
 ; CURRENT APPLICATION NUMBER: US/09/004,406C  
 ; CURRENT FILING DATE: 1998-01-08  
 ; PRIOR APPLICATION NUMBER: 08/637,640  
 ; PRIOR FILING DATE: 1996-08-23  
 ; PRIOR APPLICATION NUMBER: 08/148,132  
 ; PRIOR FILING DATE: 1993-11-02  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 3038  
 ; TYPE: PRT  
 ; ORGANISM: TPKS Protein  
 US-09-004-406C-2

Query Match Best Local Similarity 4.3%; Score 87; DB 4; Length 3038;  
 Matches 93; Conservative 55; Mismatches 138; Indels 186; Gaps 24;

QY 6 IAFCTTLALSYSPVDNGVEGEPIEC--GPTSI-----TINENT 44  
 DB 1175 VAFQVIGV--YSSP-----GDRRLRCLYPTVHVDRLIYPSLCLATAESGCEKVAFT 1226  
 QY 45 RIAFGHYVVG--LYDEGCRNDEGGQVAGISL--PPDSCNVAKTSLSNRRGIF--VT 98  
 DB 1227 INTYKRGYLSGDIVVPAE---QTLFQYENITFKFPSPDASTDHAFRMSWGPLET 1282  
 QY 99 TTVVLSFHLFETKYDR-----AVRQCFYMEADKTVSAQIEVSEITTA---FOTOIYPM 150  
 DB 1283 PDSLNDNEYMATADCKEAIPIETRIYFTI--RSFLSQLTLERQOAHFLQKQIEWL 1339  
 QY 151 PVCRYEILDGPTGPVQFAIGQPVYHKWTCSETVDTFCAVHSCFVDD-----201  
 DB 1340 E-----QVLASAKGR-----HLWYDPGWMENDEPAQIEHLCTANSYHPRVLQ 1383  
 QY 202 -----GNGDVELIADGCAIDKYLIN-----LEYPDILAGQEAHYKYAD 244  
 DB 1394 RVGQHLLEPTVRNNGNPFLLDHDG-LTTEFYNTLSFGPALHYARELVA-QIAHRYQSM 1441  
 QY 245 RSLQFYCOISITTIKEPNSCEVRPCSEPGFCAVGTGAAKPAAQOL-----294  
 DB 1442 ILFT-----GAG---TGAGTKYVLAATPOLGNSSTYTDI 1472  
 QY 295 -----RLKRSAPENIIDVRTDITLLEISDNOALP 327  
 DB 1473 STGFPEQAREQAPPEDEHVEFPLDIRSPAEQGEF--HAYDLIIASNVLAHTPD---LE 1528  
 QY 328 VDLRH-RALLQHNQPVIL-----AAVONGICMSPF 357  
 DB 1529 KTMAMARSILKPGGQVILITHTKHTRLGIFGLFADWAGVDDGRCTEPF 1580

RESULT 12  
 US-09-134-001C-3139

; Sequence 3139, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 3139  
 ; LENGTH: 438  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-3139

Query Match Best Local Similarity 4.2%; Score 85.5; DB 4; Length 438;  
 Matches 62; Conservative 49; Mismatches 127; Indels 85; Gaps 13;

QY 32 ECGPISITINENTRNAFEG--HYVYKGLYDQSGCRNDEGGQVAGISLPPDSCNVAKTS 89  
 DB 154 ENGFTKIDPRLTASAEQSTSELFHTKYFDEDAFLSQS-----QLYMEAAAHGR- 205  
 QY 90 LNPGRGIVTTIVISFHLFETKYDRAYR--VQCFYMEADKTVSAQIEVSEITTAFOQI 147  
 DB 206 -----VFSFGPFRFAEKSTRRLHLEFWNIEPMAFTNHAESLEIOEQYVSHI 253  
 QY 148 VPMRV--CRYEILDGPTGPVQFAIGQPVYHKWTCSETVDTFCAVHSCFVDDGND 205  
 DB 254 VQSVLNHQLQELK-----ALBDDT--TKLEKVAATFPRIISYD 289  
 QY 206 TVEILNADSCALDKYLLNNLEYPTDLMAQEAHYKYADRSOLF--YQCOISITTIKEPNS 263  
 DB 290 AIEFLKKEG-----FDIEWGEDFGAPHETAIANNHYDLPVFTINYPTKIKFPYQWNP 342  
 QY 264 E-----CYRQPCSEPGFCAVGTGAAKPAAQOLRLKRSAPENIIDVRTDITL 317  
 DB 343 DNEEDTVLCA--DLAPEGGETIGGSEKIN-----DLELLEQR-----INEH 382  
 QY 318 EISDNOALPVDLRHALLQHNQ 340  
 DB 383 ELDEESYIYDLRLRYGVSYPHSG 405

## RESULT 13

US-08-682-517-15  
 ; Sequence 15, Application US/08682517  
 ; Patent No. 5874267  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: Expression of surface layer proteins  
 ; NUMBER OF SEQUENCES: 25  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/682,517  
 ; FILING DATE:

;; CLASSIFICATION:  
 ;; INFORMATION FOR SEQ ID NO: 15:  
 ;; SEQUENCE CHARACTERISTICS:  
 ;; LENGTH: 1222 amino acids  
 ;; TYPE: amino acid  
 ;; TOPOLOGY: linear  
 ;; MOLECULE TYPE: protein  
 US-08-682-517-15









FILING DATE: 27-JUL-1996  
APPLICATION NUMBER: 9622617.0  
FILING DATE: 30-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31354-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-785-076-2

Query Match 4.0%; Score 80.5; DB 1; Length 432;  
Best Local Similarity 19.4%; Pred. No. 1.4;  
Matches 55; Conservative 46; Mismatches 103; Indels 79; Gaps 14;

OY 102 VISFPLEVTKVDRAYR--VOCFVMEADKTVSAQIEVSEITTAFOQTIVP--MPVCRYEI 157  
| | | | | : : : : : | | | | : : : : :  
DB 200 VFSFGPTFRAEKSRTKRLIEFMIEGMAFTNHAESLEIOQYTVHYKSVLENCLEL 259  
| | | | | : : : : : | | | | : : : : :  
OY 158 LDGPTGQPVQFAIIGQVYHKWTCDSFTVDT--FCVVHSCFVDDGNDTVEILNADCA 216  
| | | | | : : : : : | | | | : : : : :  
DB 260 -----KILERDTSKLEKVAHPFRISYD-----DAIEFLKAE-- 292  
| | | | | : : : : : | | | | : : : : :  
OY 217 LDKYLLNNLEYPITDLMAQGEAHVYKADRSOLF--YQCIISITIKENPSECVRP--QCSE- 272  
| | | | | : : : : : | | | | : : : : :  
DB 293 -----FDIEWGEDEFGAHERTALIANHYDLPVFTNYPTRKIPFYQPNPENEETVLCADL 347  
| | | | | : : : : : | | | | : : : : :  
OY 273 --POGFGAVKTGGAAPAAQAOLRLKRSAPENIIDVKTIDNTLEISDNOALPYDL 330  
| | | | | : : : : : | | | | : : : : :  
DB 348 IAPGTYGRI-IGGS-----ERVDLELLERQVKEH-----GLDEBAVSYIYDL 389  
| | | | | : : : : : | | | | : : : : :  
OY 331 RHRALLHNGOPVILAAVONGICMSPFGFSFMGISTALIAV 373  
| | | | | : : : : : | | | | : : : : :  
DB 390 RRYGSVPH-----C-----GFGLELERTVAMISGV 414  
| | | | | : : : : : | | | | : : : : :

RESULT 22  
US-09-018-824-2  
Sequence 2, Application US/09018824  
Patent No. 6410286  
GENERAL INFORMATION:  
APPLICANT: Hodgson, John  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NO. 6410286el tRNA Synthetase  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smltkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,824  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/785,076  
FILING DATE: 17-JAN-1997

APPLICATION NUMBER: 9601096.2  
FILING DATE: 19-JAN-1996  
APPLICATION NUMBER: 9615845.6  
FILING DATE: 27-JUL-1996  
APPLICATION NUMBER: 9622617.0  
FILING DATE: 30-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31354-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-018-824-2

Query Match 4.0%; Score 80.5; DB 4; Length 432;  
Best Local Similarity 19.4%; Pred. No. 1.4;  
Matches 55; Conservative 46; Mismatches 103; Indels 79; Gaps 14;

OY 102 VISFPLEVTKVDRAYR--VOCFVMEADKTVSAQIEVSEITTAFOQTIVP--MPVCRYEI 157  
| | | | | : : : : : | | | | : : : : :  
DB 200 VFSFGPTFRAEKSRTKRLIEFMIEGMAFTNHAESLEIOQYTVHYKSVLENCLEL 259  
| | | | | : : : : : | | | | : : : : :  
OY 158 LDGPTGQPVQFAIIGQVYHKWTCDSFTVDT--FCVVHSCFVDDGNDTVEILNADCA 216  
| | | | | : : : : : | | | | : : : : :  
DB 260 -----KILERDTSKLEKVAHPFRISYD-----DAIEFLKAE-- 292  
| | | | | : : : : : | | | | : : : : :  
OY 217 LDKYLLNNLEYPITDLMAQGEAHVYKADRSOLF--YQCIISITIKENPSECVRP--QCSE- 272  
| | | | | : : : : : | | | | : : : : :  
DB 293 -----FDIEWGEDEFGAHERTALIANHYDLPVFTNYPTRKIPFYQPNPENEETVLCADL 347  
| | | | | : : : : : | | | | : : : : :  
OY 273 --POGFGAVKTGGAAPAAQAOLRLKRSAPENIIDVKTIDNTLEISDNOALPYDL 330  
| | | | | : : : : : | | | | : : : : :  
DB 348 IAPGTYGRI-IGGS-----ERVDLELLERQVKEH-----GLDEBAVSYIYDL 389  
| | | | | : : : : : | | | | : : : : :  
OY 331 RHRALLHNGOPVILAAVONGICMSPFGFSFMGISTALIAV 373  
| | | | | : : : : : | | | | : : : : :  
DB 390 RRYGSVPH-----C-----GFGLELERTVAMISGV 414  
| | | | | : : : : : | | | | : : : : :

RESULT 23  
US-08-484-575A-20  
Sequence 20, Application US/08484575A  
Patent No. 5925358  
GENERAL INFORMATION:  
APPLICANT: Mark D. Cochran and David E. Junker  
TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,575A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:



```

; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0450
; TELEFAX: (212)391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-575A-20

Query Match          3.9%; Score 78.5; DB 2; Length 434;
Best Local Similarity 18.9%; Pred. No.2.5; Mismatches 148; Indels 141; Gaps 22;
Matches 80; Conservative 55;

QY 45 RNAPEGHVYKGLYDEGCNDEGROVAGISLPFSCNVAARTSLNPGIFVTTTVIS 104
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 55 RNAMDRHLFLRNAFWTIVLSSPASOSTAVTYD-----ILGRALDALTIPAVG 105
QY 105 FHLEFVTKVDRAVRVOCFYME-----ADKTVSAQIEVSEITTAFTQTIVPMPCRYEIL 158
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 106 PYNRYLITRVSRG-----CDVVELNPISNVDMMISAKEKEK----- 141
QY 159 DGGP-TGQPYQFAII-----GQ-----PYRHK--WTCSEIVDFPCAV-----VHSC 197
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 -GGPFASVYWFYVIRKGDDEDEKCPYRKYREKCGDVQLSSCAVQSAQMAVDPVPS 200
QY 198 FVDDGNGDVEIILNAGCALDKYLL----- 222
DB 201 LV-SRNGAGLITFSPALASGYLLTLKIGRFAQTALVLEVNDRLKIGSQLNPLPSKC 259
QY 223 -NNLEYPTDLMAQEAHVYKYADRS-----OLFYOQISITTIKEPSECVRPQ 269
DB 260 WTEEQYQTFG--QGSHLPFIADTNRHADVYRGYEDILQRRNNILRRKNPSAPDRPD 316
QY 270 CSEPGGFAVKTGGAAPAAQAQLRLKRSAPENIIDVPTDINTLEISDNQA---- 325
DB 317 -SVQGEIPAVTKKAEGRTDAES-----SEKKAPPE--DSEDDMQA-EASGENPALPE 366
QY 326 ---LPVDLHRRLAQ----HNGQPYLLAAVQ-----NGICMSFGFSMGISIALIAAV 373
DB 367 DDEVPEDTEHDDPNSDPDYNDMPAVIPVEETTKSSNAVSM-PI-FAAFVACAVALLVGLL 424
QY 374 IITI 377
DB 425 VMSI 428

RESULT 24
US-08-477-459-20
; Sequence 20, Application US/08477459
; Patent No. 6001369
; GENERAL INFORMATION:
; APPLICANT: Mark D. Cochran
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,459
; FILING DATE: 07-JUN-1995
```

```

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-459-20

Query Match          3.9%; Score 78.5; DB 3; Length 434;
Best Local Similarity 18.9%; Pred. No.2.5; Mismatches 148; Indels 141; Gaps 22;
Matches 80; Conservative 55;

QY 45 RNAPEGHVYKGLYDEGCNDEGROVAGISLPFSCNVAARTSLNPGIFVTTTVIS 104
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 55 RNAMDRHLFLRNAFWTIVLSSPASOSTAVTYD-----ILGRALDALTIPAVG 105
QY 105 FHLEFVTKVDRAVRVOCFYME-----ADKTVSAQIEVSEITTAFTQTIVPMPCRYEIL 158
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 106 PYNRYLITRVSRG-----CDVVELNPISNVDMMISAKEKEK----- 141
QY 159 DGGP-TGQPYQFAII-----GQ-----PYRHK--WTCSEIVDFPCAV-----VHSC 197
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 -GGPFASVYWFYVIRKGDDEDEKCPYRKYREKCGDVQLSSCAVQSAQMAVDPVPS 200
QY 198 FVDDGNGDVEIILNAGCALDKYLL----- 222
DB 201 LV-SRNGAGLITFSPALASGYLLTLKIGRFAQTALVLEVNDRLKIGSQLNPLPSKC 259
QY 223 -NNLEYPTDLMAQEAHVYKYADRS-----OLFYOQISITTIKEPSECVRPQ 269
DB 260 WTEEQYQTFG--QGSHLPFIADTNRHADVYRGYEDILQRRNNILRRKNPSAPDRPD 316
QY 270 CSEPGGFAVKTGGAAPAAQAQLRLKRSAPENIIDVPTDINTLEISDNQA---- 325
DB 317 -SVQGEIPAVTKKAEGRTDAES-----SEKKAPPE--DSEDDMQA-EASGENPALPE 366
QY 326 ---LPVDLHRRLAQ----HNGQPYLLAAVQ-----NGICMSFGFSMGISIALIAAV 373
DB 367 DDEVPEDTEHDDPNSDPDYNDMPAVIPVEETTKSSNAVSM-PI-FAAFVACAVALLVGLL 424
QY 374 IITI 377
DB 425 VMSI 428

RESULT 25
US-08-479-869-20
; Sequence 20, Application US/08479869
; Patent No. 6123949
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D, Mark D
; TITLE OF INVENTION: Recombinant Fowlpox Virus S-PPV-043 and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```





```

; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-11

```

```

Query Match          3.9%; Score 78.5; DB 5; Length 434;
Best Local Similarity 18.9%; Pred. No. 2.5; Mismatches 148; Indels 141; Gaps 22;
Matches 80; Conservative 55;

```

```

QY 45 RNAFECHVYVYKGLYDQECRNDEGGROVAGISLPFDSQNVARTSLNPRGIFVTTTVIS 104
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 RNAMDRHLFLRNAFWITVLSSFASQSTAAVTYD-----ILGRALDALTIPIAVG 105

QY 105 FHPLFVTKVDRAYVQCFYME-----ADKTVSAQIEVSEITTAFTQTQIYPMPCRYEIL 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 PYNNYLTIRVSRG---CDVVELNPISNVDDMISAKEKER-----141

QY 159 DGGP-TGQPVQFALIT---GQ---PYVHK--WTCDSFTVDTFCAY-----VHSC 197
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 -GGPFASVWFYVYIKGDDEGDKTCPIYRKREYRCGQVQLSECAVSAQMAVAVYVST 200

QY 198 FVDDNGDVEILNADGALDKYLL-----222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 LV-SRNAGGLTFSPALSGQYLLTLKIGFAQTALVLEVNDRCIKISQLNPLPSKC 259

QY 223 -NNLEPTDLMAQGEAHVYKADRS-----QLEYQCOISITIKEPNSECYRQ 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 WTEQYOTGF---QGEHLVYPIADTNTRHADVYRGYEDILQRMNNLLRKKNPSPADPRPD 316

QY 270 CSEPOGGAVKTGGAAPAAQALRLKKRSAPENIIDVRTDINTLEISDDNQ----- 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 317 -SVPOEIPAVYTKKAGRTPDAS-----SEKKAPPE--DSEDMDQA-EASGENPALPE 366

QY 326 ---LPVDLRHRLIQ---HNGQPVIIAAVQ---NGICSPFGFSFMGLSIALIAAV 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 367 DDEVPEDETHDPPNSDDPYNDMPAVIPVEETTKSSNAVSM-PI-FAAFVACAVALVGLL 424

QY 374 IITIT 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 VMSI 428

```

```

RESULT 30
PCT-US96-03916-67
; Sequence 67, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-67

```

```

Query Match          3.9%; Score 78.5; DB 5; Length 434;
Best Local Similarity 18.9%; Pred. No. 2.5; Mismatches 148; Indels 141; Gaps 22;
Matches 80; Conservative 55;

```

```

QY 45 RNAFECHVYVYKGLYDQECRNDEGGROVAGISLPFDSQNVARTSLNPRGIFVTTTVIS 104
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 RNAMDRHLFLRNAFWITVLSSFASQSTAAVTYD-----ILGRALDALTIPIAVG 105

QY 105 FHPLFVTKVDRAYVQCFYME-----ADKTVSAQIEVSEITTAFTQTQIYPMPCRYEIL 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 PYNNYLTIRVSRG---CDVVELNPISNVDDMISAKEKER-----141

QY 159 DGGP-TGQPVQFALIT---GQ---PYVHK--WTCDSFTVDTFCAY-----VHSC 197
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 -GGPFASVWFYVYIKGDDEGDKTCPIYRKREYRCGQVQLSECAVSAQMAVAVYVST 200

QY 198 FVDDNGDVEILNADGALDKYLL-----222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 LV-SRNAGGLTFSPALSGQYLLTLKIGFAQTALVLEVNDRCIKISQLNPLPSKC 259

QY 223 -NNLEPTDLMAQGEAHVYKADRS-----QLEYQCOISITIKEPNSECYRQ 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 WTEQYOTGF---QGEHLVYPIADTNTRHADVYRGYEDILQRMNNLLRKKNPSPADPRPD 316

QY 270 CSEPOGGAVKTGGAAPAAQALRLKKRSAPENIIDVRTDINTLEISDDNQ----- 325
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 317 -SVPOEIPAVYTKKAGRTPDAS-----SEKKAPPE--DSEDMDQA-EASGENPALPE 366

QY 326 ---LPVDLRHRLIQ---HNGQPVIIAAVQ---NGICSPFGFSFMGLSIALIAAV 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 367 DDEVPEDETHDPPNSDDPYNDMPAVIPVEETTKSSNAVSM-PI-FAAFVACAVALVGLL 424

QY 374 IITIT 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 VMSI 428

```

```

RESULT 31
US-08-484-993B-16
; Sequence 16, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunoreception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-993B-16

Query Match      3.9%; Score 78.5; DB 2; Length 570;
Best Local Similarity 21.1%; Pred. No. 4;
Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

OY 17 YSIPVNGVEGPEIECGPTSTITNPNTRNAFEGHYVYKGLYDQCGRNDEGGROYAGIS 76
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 230 FSIANSRVNTSP-----ILLN-SLRLAFG-----KDR-CNPVATRAFALEF 271
OY 77 LPFDSGNVARTSLNPRGIFVTYTVVISFHPLEVTYKVDRA-----RVOCFY 123
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 272 FPFNSCGTTR-----WYTGDAVYENELVARDVRTWSHGSTRISIRLAVSCSY 322
OY 124 MEADTVSAQILEVSEITTA-FOTQIVMPV-----CRYEILDGPTGQPVQPAI 171
DB 323 SVRSNAFPLSVQVFTIPRPHLKTGHGPLTLELKIADKHGSGYYTIGD-----YFV-VKL 376
OY 172 IGOPYHKWTCDESEVDFCAVHSCFVDDG-NGDTVE---ILNADGALDKYLINLEY 227
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 377 LRDPIYEVSTIRHRTDPSLGLLHNCWATPKNSOSLSQMPIL-VKGC---PYVDN--Y 430
OY 228 PTDLMAQGEA-----HYVYADRSQLFYOCQISITIKEPNSQC 265
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 431 QTQLIPVQKALDTPPSYKRFSTFTSFVDTPMAKMLRGVYLHCNVSICQPACTSSC 489

RESULT 32
US-08-484-158B-16
; Sequence 16, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunoncontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
```

```

;
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-158B-16

Query Match      3.9%; Score 78.5; DB 2; Length 570;
Best Local Similarity 21.1%; Pred. No. 4;
Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

OY 17 YSIPVNGVEGPEIECGPTSTITNPNTRNAFEGHYVYKGLYDQCGRNDEGGROYAGIS 76
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 230 FSIANSRVNTSP-----ILLN-SLRLAFG-----KDR-CNPVATRAFALEF 271
OY 77 LPFDSGNVARTSLNPRGIFVTYTVVISFHPLEVTYKVDRA-----RVOCFY 123
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 272 FPFNSCGTTR-----WYTGDAVYENELVARDVRTWSHGSTRISIRLAVSCSY 322
OY 124 MEADTVSAQILEVSEITTA-FOTQIVMPV-----CRYEILDGPTGQPVQPAI 171
DB 323 SVRSNAFPLSVQVFTIPRPHLKTGHGPLTLELKIADKHGSGYYTIGD-----YFV-VKL 376
OY 172 IGOPYHKWTCDESEVDFCAVHSCFVDDG-NGDTVE---ILNADGALDKYLINLEY 227
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 377 LRDPIYEVSTIRHRTDPSLGLLHNCWATPKNSOSLSQMPIL-VKGC---PYVDN--Y 430
OY 228 PTDLMAQGEA-----HYVYADRSQLFYOCQISITIKEPNSQC 265
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 431 QTQLIPVQKALDTPPSYKRFSTFTSFVDTPMAKMLRGVYLHCNVSICQPACTSSC 489

RESULT 33
US-08-484-596A-16
; Sequence 16, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; TITLE OF INVENTION: Materials and Methods for Immunoncontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
```

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,596A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,223  
FILING DATE: 11-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-596A-16

Query Match 3.9%; Score 78.5; DB 2; Length 570;  
Best Local Similarity 21.1%; Pred. No. 4;

Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

QY 17 YSIIPVNGVEGEPEIECGPSTITINFTNRNAFEGHVVYKGLYDGGGRNDGGRQVAGIS 76  
DB 230 FSIIVSRNVTSP-----LLLN-SRLAFG-----KDRE-CNPVKATRAFLFF 271  
QY 77 LRPDSCNVAARTSLNPGIEVTTTVISFHPLEVTKYDRAY-----RVOCFY 123  
DB 272 FPNNSCGTTR-----WVTGDAVYENELVAARDVRWMSHSITRDSIFRLRVGCSY 322  
QY 124 MEADKTVSAQIEVSEITTA-FOTOIVPMV-----CRYEILDGPGTGPVOPAI 171  
DB 323 SVRSNAFPLSVQVFTTIPPHLKTQHGPLELTKAKDKHGYSTYIGD-----YPV-VKL 376  
QY 172 IGPVYHKMTCDESETVDFCAVYVHSCFVDDG-NGDIVE---ILNADGALDKYLLNMLEY 227  
DB 377 LRDPYIEVSIIRHTDPSLGLLHNCWATPGKNSQSLSQWPI-L-VKGC---PYVGDN--Y 430  
QY 228 PTDLMAQEA-----HYKKYADRSOLFYOQOISITIKPENSEC 265  
DB 431 QTOILPVQKALDTPFPYSYKRFSTFTSFYDTMAKMLRGFVYLHCNVISICOPAGTSSC 489

RESULT 34

US-08-480-150A-16  
Sequence 16, Application US/08480150A  
Patent No. 598550  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,150A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223  
FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-480-150A-16

Query Match 3.9%; Score 78.5; DB 2; Length 570;  
Best Local Similarity 21.1%; Pred. No. 4;

Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

QY 17 YSIIPVNGVEGEPEIECGPSTITINFTNRNAFEGHVVYKGLYDGGGRNDGGRQVAGIS 76  
DB 230 FSIIVSRNVTSP-----LLLN-SRLAFG-----KDRE-CNPVKATRAFLFF 271  
QY 77 LRPDSCNVAARTSLNPGIEVTTTVISFHPLEVTKYDRAY-----RVOCFY 123  
DB 272 FPNNSCGTTR-----WVTGDAVYENELVAARDVRWMSHSITRDSIFRLRVGCSY 322  
QY 124 MEADKTVSAQIEVSEITTA-FOTOIVPMV-----CRYEILDGPGTGPVOPAI 171  
DB 323 SVRSNAFPLSVQVFTTIPPHLKTQHGPLELTKAKDKHGYSTYIGD-----YPV-VKL 376  
QY 172 IGPVYHKMTCDESETVDFCAVYVHSCFVDDG-NGDIVE---ILNADGALDKYLLNMLEY 227  
DB 377 LRDPYIEVSIIRHTDPSLGLLHNCWATPGKNSQSLSQWPI-L-VKGC---PYVGDN--Y 430  
QY 228 PTDLMAQEA-----HYKKYADRSOLFYOQOISITIKPENSEC 265  
DB 431 QTOILPVQKALDTPFPYSYKRFSTFTSFYDTMAKMLRGFVYLHCNVISICOPAGTSSC 489

RESULT 35

US-08-458-731-16  
Sequence 16, Application US/08458731  
Patent No. 6001599  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59

```

CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 570 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-731-16

Query Match
Best Local Similarity 21.1%; Score 78.5; DB 3; Length 570;
Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

QY 17 YSLPVDNGEGPEIECGTSTITNNTNAFEGHYVYKGLYDQEGCRNDEGGROYAGIS 76
DB 230 FSIASRVNTSP-----LLLN-SRLAFG-----KDRE-CNPVKATRAFALEF 271
QY 77 LPFDSGNVARTSLNPRGIFVTTVVISFHPLEVTKVDRA-----RVQCFY 123
DB 272 PPFNSCGTTR-----WYTGDAVYENELVAARDVFTWSHGSIITRDSIFRLRVSCSY 322
QY 124 MEADTVSAQIEVSEITTA-FQTOIVPMV-----CRYEILDGPGTGPVQFAI 171
DB 323 SVRSNAFPLSVQVFTIPRPHLKTQHGRLLELKIADKHGYSYITGD-----YV-VKL 376
QY 172 IGQPVYHKWTCDSFVDFPCAVVHSCFVDDG-NGDTEV---ILNADGALDKYLLINLEY 227
DB 377 LRDPIYEVVISIRHRTDPSLIGLLHNCMATPGKNSQSLQWPII-VKGC---PYVGDN--Y 430
QY 228 PTDLMAQGEA-----HYVYADRSLQFYQCOISITIKENNSC 265
DB 431 QTOILPVQKALDTPPSPYTKRFSTFTSFVDPMKMKMLRGVYLHCNVISICOPAGTSSC 489

RESULT 36
US-08-149-223A-16
Sequence 16, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 570 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-149-223A-16

Query Match
Best Local Similarity 21.1%; Score 78.5; DB 3; Length 570;
Matches 63; Conservative 38; Mismatches 109; Indels 89; Gaps 16;

QY 17 YSLPVDNGEGPEIECGTSTITNNTNAFEGHYVYKGLYDQEGCRNDEGGROYAGIS 76
DB 230 FSIASRVNTSP-----LLLN-SRLAFG-----KDRE-CNPVKATRAFALEF 271
QY 77 LPFDSGNVARTSLNPRGIFVTTVVISFHPLEVTKVDRA-----RVQCFY 123
DB 272 PPFNSCGTTR-----WYTGDAVYENELVAARDVFTWSHGSIITRDSIFRLRVSCSY 322
QY 124 MEADTVSAQIEVSEITTA-FQTOIVPMV-----CRYEILDGPGTGPVQFAI 171
DB 323 SVRSNAFPLSVQVFTIPRPHLKTQHGRLLELKIADKHGYSYITGD-----YV-VKL 376
QY 172 IGQPVYHKWTCDSFVDFPCAVVHSCFVDDG-NGDTEV---ILNADGALDKYLLINLEY 227
DB 377 LRDPIYEVVISIRHRTDPSLIGLLHNCMATPGKNSQSLQWPII-VKGC---PYVGDN--Y 430
QY 228 PTDLMAQGEA-----HYVYADRSLQFYQCOISITIKENNSC 265
DB 431 QTOILPVQKALDTPPSPYTKRFSTFTSFVDPMKMKMLRGVYLHCNVISICOPAGTSSC 489

RESULT 37
US-08-325-071-61
Sequence 61, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold

```



APPLICANT: SRISKANTHA, Alagaccone  
APPLICANT: RIDING, George Alfred  
APPLICANT: RAND, Keith No. 5587311man  
TITLE OF INVENTION: DNA Encoding A Cell Membrane  
TITLE OF INVENTION: Glycoprotein Of A Tick Gut  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/325,071  
FILING DATE: 14-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/062,109  
FILING DATE: 17-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,368  
FILING DATE: 07-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/242,196  
FILING DATE: 06-JUL-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU87/00401  
FILING DATE: 27-NOV-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P14912  
FILING DATE: 16-OCT-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P12570  
FILING DATE: 19-JUN-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P9196  
FILING DATE: 27-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
TELEPHONE/DOCKET NUMBER: 60042/111 BIAU  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 549 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-325-071-61

Query Match 3.88; Score 78; DB 1; Length 549;  
Best Local Similarity 23.48; Pred. NO. 4.2;  
Matches 62; Conservative 26; Mismatches 95; Indels 82; Gaps 15;

QY 115 RAYVOC---FYMEADKTVSAQIEVSEITTAFOFQIV-PMVCHYE--IIDSGPTGQPV 167  
Db 146 KAYCTCPRGTVYEDGTTCSISHTVSCIAEQKOTCPFTDCRVHKGTVLCECPWNO-- 203  
QY 168 QFALIGQPVYHAKWTCDSEVTDTFCVAVHSCFVDDG-----NGDPTVELIN 211  
Db 204 --HLVGD-----TCISQCVDKC--HEEFMDGQVWNRQSCYCPMKSRKRGPNV--N 249  
QY 212 ADGCAIDKYLIN-----NLEPPTLMAGQEAHYKTVADRSQLFYQCOISTIKKEPNSECV 266  
Db 250 INGLINEYYTVSFTPNISPSD-----HCKWYEDR--VLEAIRTSIGKEVFKYEIL 300

QY 267 RPOCSEPOGFCGAVKTCGAARPAAPAAOLR-----LLKRSR--EP 304  
Db 301 --NCTQD-----IKARLAEKPLSNHVLRKIQACEHPIGEMOMYPKLLIKNSATEIEE 353  
QY 305 ENIIDVTDITLLEISDNOALPYD 329  
Db 354 ENLCDSLKNOEAAYKGNCKVKVD 378

RESULT 38  
US-08-461-004A-61  
Sequence 61, Application US/08461004A  
Patent No. 6235283  
GENERAL INFORMATION:  
APPLICANT: COBON, Stewart Gary  
APPLICANT: MOORE, Joanna Terry  
APPLICANT: JOHNSON, Law Anthony York  
APPLICANT: WILADSEN, Peter  
APPLICANT: KEMP, David Harold  
APPLICANT: SRISKANTHA, Alagaccone  
APPLICANT: RIDING, George Alfred  
APPLICANT: RAND, Keith No. 6235283man  
TITLE OF INVENTION: DNA Encoding A Cell Membrane  
TITLE OF INVENTION: Glycoprotein Of A Tick Gut  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,004A  
FILING DATE: 04-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/325,071  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/062,109  
FILING DATE: 17-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,368  
FILING DATE: 07-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/242,196  
FILING DATE: 06-JUL-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU87/00401  
FILING DATE: 27-NOV-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P14912  
FILING DATE: 16-OCT-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P12570  
FILING DATE: 19-JUN-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P9196  
FILING DATE: 27-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
TELEPHONE/DOCKET NUMBER: 60042/152  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:





**THIS PAGE BLANK (USPTO)**

Gencore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 24, 2003, 08:52:19 ; Search time 18 seconds  
(without alignments)  
1149.438 Million cell updates/sec

Title: US-10-054-562A-4

Perfect score: 2026

Sequence: 1 MMIRLAFCTTLALSYSTIP.....ALIAAYITISFKRENOKA 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

```
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2026	100.0	387	9	US-10-054-562A-4
2	1086.5	53.6	245	9	US-10-054-562A-17
3	776.5	38.3	271	9	US-10-054-562A-9
4	92	4.5	896	10	US-09-923-563A-1
5	87.5	4.3	474	10	US-09-833-147-28
6	87	4.3	473	10	US-09-833-147-29
7	85.5	4.2	1118	9	US-09-909-567B-42
8	85	4.2	2069	9	US-09-738-626-4320
9	84.5	4.2	475	10	US-09-729-454-1
10	83	4.1	1222	10	US-09-137-531-15
11	83	4.1	1252	10	US-09-137-531-9
12	81.5	4.0	1948	9	US-09-808-602-55
13	80	3.9	293	9	US-09-738-626-5474
14	80	3.9	26926	9	US-09-759-508B-2
15	79.5	3.9	478	10	US-09-835-147-8
16	78.5	3.9	299	9	US-09-974-298-132
17	78.5	3.9	299	10	US-09-799-848-1
18	78.5	3.9	434	10	US-09-881-457A-4
19	77.5	3.8	646	9	US-10-121-746-10

20	77	3.8	1706	10	US-09-864-761-46862	Sequence 46862, A
21	77	3.8	4563	9	US-09-870-759-128	Sequence 128, App
22	77	3.8	4563	9	US-09-802-640-92	Sequence 32, Appl
23	76.5	3.8	295	9	US-09-738-626-4782	Sequence 4782, App
24	76.5	3.8	1184	9	US-09-992-598-124	Sequence 124, App
25	76.5	3.8	1184	9	US-09-989-293A-124	Sequence 124, App
26	76.5	3.8	1184	9	US-09-989-735-124	Sequence 124, App
27	76.5	3.8	1184	9	US-09-980-444-124	Sequence 124, App
28	76.5	3.8	1184	9	US-09-989-730-124	Sequence 124, App
29	76.5	3.8	1184	9	US-09-990-436-124	Sequence 124, App
30	76.5	3.8	1184	9	US-09-991-181-124	Sequence 124, App
31	76.5	3.8	1184	9	US-09-993-687-124	Sequence 124, App
32	76.5	3.8	1184	9	US-09-989-734-124	Sequence 124, App
33	76.5	3.8	1184	9	US-10-028-072-412	Sequence 412, App
34	76.5	3.8	1184	9	US-09-997-653-124	Sequence 124, App
35	76.5	3.8	1184	9	US-09-993-667-124	Sequence 124, App
36	76.5	3.8	1184	9	US-10-121-049-412	Sequence 412, App
37	76.5	3.8	1184	9	US-10-123-904-412	Sequence 412, App
38	76.5	3.8	1184	9	US-10-140-470-412	Sequence 412, App
39	76.5	3.8	1184	9	US-09-990-438-124	Sequence 124, App
40	76.5	3.8	1184	9	US-09-990-562-124	Sequence 124, App
41	76.5	3.8	1184	9	US-09-997-428-124	Sequence 124, App
42	76.5	3.8	1184	9	US-09-997-666-124	Sequence 124, App
43	76.5	3.8	1184	9	US-10-175-746-412	Sequence 412, App
44	76.5	3.8	1184	9	US-10-176-918-412	Sequence 412, App
45	76.5	3.8	1184	9	US-10-176-921-412	Sequence 412, App

## ALIGNMENTS

```
RESULT 1
US-10-054-562A-4
; Sequence 4, Application US/10054562A
; Patent No. US20020165375A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekhar, Ramaswamy
; TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES,
; FILE REFERENCE: HW-8-2
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-10-054-562A-4

Query Match      100.0%; Score 2026; DB 9; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.1e-201; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 0;

QY 1 MMIRLAFCTTLALSYSTIPVDNGVEGEPEETCTSTTINFNRNAFEGHVYKGLYDQ 60
1 MMIRLAFCTTLALSYSTIPVDNGVEGEPEETCTSTTINFNRNAFEGHVYKGLYDQ 60
Db 1 MMIRLAFCTTLALSYSTIPVDNGVEGEPEETCTSTTINFNRNAFEGHVYKGLYDQ 60
QY 61 EGCNNDGGROVAGISLPDSCNVARFRSLNPRGIFVTTTVVISFHPLEVTKVDRAVRVQ 120
61 EGCNNDGGROVAGISLPDSCNVARFRSLNPRGIFVTTTVVISFHPLEVTKVDRAVRVQ 120
Db 61 EGCNNDGGROVAGISLPDSCNVARFRSLNPRGIFVTTTVVISFHPLEVTKVDRAVRVQ 120
QY 121 CFVWEAKTVSAQLEVESEITAFQTOIVPMVCXREILIDGPTGQPVQAFATIGPVRVHKW 180
121 CFVWEAKTVSAQLEVESEITAFQTOIVPMVCXREILIDGPTGQPVQAFATIGPVRVHKW 180
Db 121 CFVWEAKTVSAQLEVESEITAFQTOIVPMVCXREILIDGPTGQPVQAFATIGPVRVHKW 180
QY 181 TCDEETVDFCAVAVHSCFVDDGNGDTVEILNADSCALDKYLLNLEKPTDLMAGOEAHVY 240
181 TCDEETVDFCAVAVHSCFVDDGNGDTVEILNADSCALDKYLLNLEKPTDLMAGOEAHVY 240
```

```
DB 181 TCSEVDFPCAVVHSCFVDDGNGDVEILNADGALDKYLNNLEYPDTLMAGQEAHY 240
;
QY 241 KYADRSOLFQYQOISITIKENPSECVAPQCSFPGFGAVTGAAGPAAAOQLRLKKR 300
;
DB 241 KYADRSOLFQYQOISITIKENPSECVAPQCSFPGFGAVTGAAGPAAAOQLRLKKR 300
;
QY 301 SAEPEINIDVTDINTLEISDDNOALPYDHRHALLQHNQPTILAVONGICMSPFGS 360
;
DB 301 SAEPEINIDVTDINTLEISDDNOALPYDHRHALLQHNQPTILAVONGICMSPFGS 360
;
QY 361 MEMGLSIALIAVITITSEKFRPNOKA 387
;
DB 361 MEMGLSIALIAVITITSEKFRPNOKA 387
;
RESULT 2
US-10-054-562a-17
; Sequence 17, Application US/10054562A
; Patent No. US20020165375A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HM-8-2
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Brugia malayi
US-10-054-562a-17
;
Query Match 53.6%; Score 1086.5; DB 9; Length 245;
Best Local Similarity 80.6%; Pred. No. 1,4e-104;
Matches 195; Conservative 24; Mismatches 22; Indels 1; Gaps 1;
;
QY 3 IRILACTLLIALST-SIVVDNGVEGEPEIEGPTSIITNENRNFEGHYVYKGLYDQ 61
;
DB 4 MOICSLSTMIASINAIIDNGVESEPEIEGPTSIITNENRNFEGHYVYKGLYDQ 63
;
QY 62 GCRNDEGGRQVAGISLPFSCNVAARTSRINPRTGIFVTTVVISFHPDLFTYKVDRAVYQC 121
;
DB 64 DCRSDGGRQVAGISLPFSCNVAARTSRINPRTGIFVTTVVISFHPDLFTYKVDRAVYQC 123
;
QY 122 FYMEADKTVSAOIEVSEITTAFTQIVPMPVCRYEILIDGPGTGOVQFAIIGQPVYHKMT 181
;
DB 124 FYMEADKTVSTQIEVSEMTTVAFTQIVPMPVCRYEILIDGPGTGOVQFAIIGQPVYHKMT 183
;
QY 182 CSEVDFPCAVVHSCFVDDGNGDVEILNADGALDKYLNNLEYPDTLMAGQEAHY 241
;
DB 184 CSEVDFPCAVVHSCFVDDGNGDVEILNADGALDKYLNNLEYPDTLMAGQEAHY 243
;
QY 242 YA 243
;
DB 244 YA 245
;
RESULT 3
US-10-054-562a-9
; Sequence 9, Application US/10054562A
; Patent No. US20020165375A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekar, Ramaswamy
; APPLICANT: Morales, Tony H.
; TITLE OF INVENTION: PARASITIC HELMINTH CUTICLIN PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HM-8-2
```

```
;
; FILE REFERENCE: HM-8-2
; CURRENT APPLICATION NUMBER: US/10/054,562A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/812,642
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 09/323,427
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-10-054-562a-9
;
Query Match 38.3%; Score 776.5; DB 9; Length 271;
Best Local Similarity 72.4%; Pred. No. 2.2e-72;
Matches 144; Conservative 24; Mismatches 22; Indels 9; Gaps 2;
;
QY 134 IEVSEITTAFTQIVPMPVCRYEILIDGPGTGOVQFAIIGQPVYHKMTDSEVDFPCAV 193
;
DB 15 LEVSEMTTAFQIQVPMPCRYEILIDGPGTGOVQFAIIGQPVYHKMTDSEVDFPCAV 194
;
QY 194 VHSCEFDDGNGDVEILNADGALDKYLNNLEYPDTLMAGQEAHYKYADRSOLFQYQO 253
;
DB 75 VHSCEFDDGNGDVEILNADGALDKYLNNLEYPDTLMAGQEAHYKYADRSOLFQYQO 253
;
QY 254 ISITIKENPSECVAPQCSFPGFGAVTGAAGPAAAOQLRLKKRSAP 304
;
DB 135 ISITIKENPSECVAPQCSFPGFGAVTGAAGPAAAOQLRLKKRSAP 304
;
QY 305 ENITDVRTDINTLEISDDN 323
;
DB 195 DNTVDVSTGSGTVDTEEN 213
;
RESULT 4
US-09-923-563a-1
; Sequence 1, Application US/09923563A
; Patent No. US20020115141A1
; GENERAL INFORMATION:
; APPLICANT: Wambach, Wolfgang
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Method for Improved Production of Cyanophycin and Secondary Pr
; FILE REFERENCE: Lea 34650
; CURRENT APPLICATION NUMBER: US/09/923,563A
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Synechococcus elongatus
US-09-923-563a-1
;
Query Match 4.5%; Score 92; DB 10; Length 896;
Best Local Similarity 22.0%; Pred. No. 1.3;
Matches 91; Conservative 44; Mismatches 155; Indels 124; Gaps 22;
;
QY 36 TSLITNENRNFEGHYVYKGLYDQEGCRNDEGGRQVAGISLPFSCNVAARTSRINPRTG 95
;
DB 356 TRIEVNHDTWTLEROGTTLNTVLDPG-----ETCYLRATANLSTGCI 398
;
QY 96 FYTTTVISFHPDLFTYKVDRAVYQCIFYMEADKTVSAOIE-----VSEITTA--FQ 144
;
DB 399 AIDRDEI--HPEWNWICQARAIIIGLDIAGIDVSPDISQSLKSGVYVYVNAAPGR 456
;
QY 145 TQIVPM-----PVCRYEILIDGPGTGOVQFAIIGQPVYHKMTDSEVDFPCAVVH 195
;
DB 457 MNTNPSQGIARVAVBPVLNMLPPPGPCRPRI-FAITG-----TNGKTTYT-RLAH 506
;
QY 196 SCFVDDGNGDVEILNADGALDKYLNN-----LEYPTLMAGQEAHYKYA 243
;
```

Db	507	IC----	KQTGGVGT	YTTWDG	IYIG	YEV	EKGSD	TG	PG	SA	Q	L	I	D	P	T	----	VE	I	A	V	L	E	T	A		558																									
QY	244	D----	RSQ	L	F	Y	Q	O	I	S	I	----	T	I	K	E	----	P	N	S	E	C	R	P	O	C	S	E	P	O	G	F	A	N	K	T	G	282														
Db	559	RG	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	L	R	G	616															
QY	283	GA	A	A	P	A	A	A	Q	L	R	L	L	K	R	----	S	A	P	E	N	I	I	D	V	R	T	I	N	T	----	T	E	I	S	D	N	O	A	L	P	327										
Db	617	--	A	D	D	P	L	V	A	A	A	A	R	Q	V	A	F	Y	S	M	D	P	N	P	I	----	I	R	O	H	I	Q	O	G	G	L	A	Y	E	N	G	L	I	S	L	K	D	W	T	L	R	673
QY	328	V	D	L	R	H	R	L	L	I	H	N	G	O	P	V	L	A	----	Y	O	N	G	I	C	M	S	P	F	G	S	M	G	S	I	A	L	A	A	V	I	T	376									
Db	674	I	E	----	Q	A	E	N	V	P	I	T	L	G	A	R	A	S	P	T	I	N	A	L	A	A	----	Q	G	S	I	E	H	R	A	L	T	T	718													

```

RESULT 5
US-09-835-147-28
: Sequence 28, Application US/09835147
: Patent No. US20020002277A1
: GENERAL INFORMATION:
: APPLICANT: Maliszewski, Charles R.
: APPLICANT: Gayle III, Richard B.
: APPLICANT: Price, Virginia L.
: APPLICANT: Gimpel, Steven D.
: TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
: FILE REFERENCE: 2879-US
: CURRENT APPLICATION NUMBER: US/09/835,147
: CURRENT FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/104,585
: PRIOR FILING DATE: 1998-10-16
: PRIOR APPLICATION NUMBER: US 60/107,466
: PRIOR FILING DATE: 1998-11-06
: PRIOR APPLICATION NUMBER: US 60/149,010
: PRIOR FILING DATE: 1999-08-13
: PRIOR APPLICATION NUMBER: PCT/US99/22955
: PRIOR FILING DATE: 1999-10-13
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 28
: LENGTH: 474
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Fusion
: US-09-835-147-28

```

	Query Match	4.3%	Score 87.5:	DB 10;	Length 474;
	Best Local Similarity	20.9%;	Pred. No.1.4;		
	Matches 55;	Conservative 40;	Mismatches 75;	Indels 93;	Gaps 13.
QY	113 VDRAYRVOCFAME--ADKIVSAOIEVESEITTAFTOTIVMPVRYEELIDCGPTGPGPQFAFI	171	:	:	:
Dd	5 IDRQLLSCIALSLALATYNASASTKTQLTSTQNKALEBNKYGIYLADGSSHTSLYT--	62	:	:	:
QY	172 IGRPVYHKMTCDSETVDTFCAVH-----SCFVDDGN-----	203	:	:	:
Dd	63 -----YKMPAEKEN-DI--GVNHVCECRVKGPISKFQVNKNETIGILTCMEREAREV	113	:	:	:
QY	204 -----GDT--VEILMDGCALDKYLLNLE-----YPTD-----LMAGOEAH	238	:	:	:
Dd	114 IPRSQHGTBPVYLGLATAGMRLLRMESEELADRLVDVERLSNYPFDGARIITGGDEEG	173	:	:	:
QY	239 VY-----KVADRSQLFOCOISITIKEPNSECYVRQCSEPOGFGAVYTGGAANKP	288	:	:	:
Dd	174 AYGWITTINVLTKFSQKTRWF-----SIAPVETNNQ-----ETFGALDIGASTGY	219	:	:	:
QY	289 AAAAQRLRLKKRSAPENIIDVR	311	:	:	:
Dd	220 TFVPQNMQTE---SPDNALOFR	238	:	:	:

RESULT 6  
 US-09-835-147-29  
 Sequence 29, Application US/09835147  
 Patent No. US20020002277A1  
 GENERAL INFORMATION:  
 APPLICANT: Maliszewski, Charles R.  
 APPLICANT: Gayle III, Richard B.  
 APPLICANT: Price, Virginia L.  
 APPLICANT: Gimpel, Steven D.  
 TITLE OR INVENTION: Inhibitors of Platelet Activation and Recruitment  
 FILE REFERENCE: 2879-US  
 CURRENT APPLICATION NUMBER: US/09/835,147  
 CURRENT FILING DATE: 2001-04-13  
 PRIOR APPLICATION NUMBER: US 60/104,585  
 PRIOR FILING DATE: 1998-10-16  
 PRIOR APPLICATION NUMBER: US 60/107,466  
 PRIOR FILING DATE: 1998-11-06  
 PRIOR APPLICATION NUMBER: US 60/149,010  
 PRIOR FILING DATE: 1999-08-13  
 PRIOR APPLICATION NUMBER: PCT/US99/22955  
 PRIOR FILING DATE: 1999-10-13  
 NUMBER OF SEQ ID NOS: 31  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 29  
 LENGTH: 473  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Fusion  
 OTHER INFORMATION: construct of human CD39  
 US-09-835-147-29

```

Query Match 4.3%; Score 87; DB 10; Length 473;
Best Local Similarity 19.5%; Pred. No. 1.6;
Matches 51; Conservative 42; Mismatches 77; Indels 92; Gaps 12;

QY 113 VDRATVOCFFWEADKTVSAQIEVSETTAFOTQIVMPVCRYEILDGPGPQVQFAII 172
Db 5 IDRMOLLSICILSLALVYNSSTKKTQTLSSQNKMLPENVKYGGIIVLDGSSHTSLYI--- 61
QY 173 GQPVYHKMTGCESEVDFCAVH-----SCFVDGDN----- 203
Db 62 ----YKMPAEKEN-DT--GVVHOVEECRVKVGPGISKFEVQKVNIEIGIYLTDCMERAREVI 113
QY 204 -----GDT--VEIILADGACALDKYLLNNLE-----YPTD-----LMAGOEAHV 239
Db 114 PRSQHETPVYLGATAGRRLRLMSEELADVDVVERSLSNYPDPFGCARIIITQGEBA 173
QY 240 Y-----KYADRSOLFVQCQISITIKENSECVCVRPQCSPEQGFAGVYTGCAAAKPA 289
Db 174 YGMITINYLTKFSQTKRMF-----SIVPEETNNO-----ETFGALDLGCASTGYT 219
QY 290 AAAQLRLKKRSAPENIIDVR 311
Db 220 FVPQNMOTIE---SPDNALQFR 237

RESULT 7
US-09-909-567B-42
; Sequence 42, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DDx-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1

```



```
; SEQ ID NO 42
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-567B-42

Query Match
Best Local Similarity 25.6%; Score 85.5; DB 9; Length 1118;
Matches 52; Conservative 27; Mismatches 81; Indels 43; Gaps 12;

15 LVSIPVNDNGVEGEPEI--ECGPTSTINFTNTNAPEG-----HYVYKGLYDQSGCR 64
Db 348 LDSVYVANKAAAFKHSKSTRKYKPTPLPIKVKVKNVKEGHIKSSVSISIVSEMD---R 403
QY 65 NDEGROVAGISLPFDSN---VARTSLNPRGIF---VTTTVISFHPLEFVTKV-- 113
Db 404 SSKG--QIIGNFQAFDEDTGLPAHARYVKLEDDNMISVDSVTSSEIKLAKLPDESEHYQ 461
QY 114 DRAHYVQCFMEAD---KTVSAQ--IEVSEITTAFOQTQVPM-PVCR-----YEILD 159
Db 462 NGTYTVKIVAISEDYPRKTTGTVLINVEDINDNCPTLIEPVOTICHDAEYVNVTAEDLD 521
QY 160 GGPVGOPVOPFALIGOP--VYHKM 180
Db 522 GHNSGPFSTSVIDKPRGMAEKW 544

RESULT 8
US-09-738-626-4320
; Sequence 4320, Application US/09738626
; Publication No. US20020197605A1
GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4320
; LENGTH: 2069
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4320

Query Match
Best Local Similarity 4.2%; Score 85; DB 9; Length 2069;
Matches 73; Conservative 50; Mismatches 116; Indels 124; Gaps 17;

15 LVSIPVDN-----GVGEPEICGPTSTINFTNTNAPEGHYVYKGLYDQSGC 63
Db 1431 VEFVYVPRDDHGKSMYLNDFEGIEHLPKVELRWLNGLIGNGQRRLGCGFEVTSPLFNV--C 1488
QY 64 R-----NDEGROVAGISLPF-----DSCNVAARTSLNPRGIFVTTTVISFHPLE 109
Db 1489 RHCGHLDSEAGANSRWDHRPMCRPHRYEKEDIVYSPALGRTLKTQGVLMLLP----- 1539
QY 110 VTKVDRAYVQCFYMEADTVSAQIEVSEITTAFOQTQVPMFVYCKRIEILDGPGTGOVOP 169
```

```
Db 1540 -----EYFGSEADSKVVTSL-IAAIKIGFR-----EVLGSDP----- 1570
QY 170 AIIIGPYVHKWTDSENVDFECVAVHSCFVDDGNDGPTVELLNADGALDKYLLNLEYP 229
Db 1571 -----DHLD-----VTSVOVPRTSGDG-----ALDALLH----- 1595
QY 230 DLMAQGEAHYKYKADRSOL-----FYQCOISITIKPEPNECVARPOCSEPOGFG 277
Db 1596 DQVPGGTGYLNGFADPTKVELISRAMERYSRCCQCYDERLACP--ECLLPYRTAT--- 1650
QY 278 AVTGGGAAPAAQAOLRLKRSABENIIDVT--DINTLEISDNO-ALPVDLRHRA 334
Db 1651 LIOTSRRAAEKA---LRAILLNLSRPREETIDLSAVPDMTELEKREPENTLGSOLELRFV 1706
QY 335 LLD 337
Db 1707 MLR 1709

RESULT 9
US-09-729-454-1
; Sequence 1, Application US/09729454
; Patent No. US20020137038A1
GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 3229449CD1
US-09-729-454-1

Query Match
Best Local Similarity 4.2%; Score 84.5; DB 10; Length 475;
Matches 62; Conservative 32; Mismatches 92; Indels 143; Gaps 12;

87 TSLNPRGIFVTTTVISFHPLEFVTVDAHYVQCFYMEADTVSAQIEVSEITTAFOQT 146
Db 50 TSLFEPK-----TPLISLKPLETE-----ELRIKETIEKLDQ 82
QY 147 IYPMVPCRYEILDGPGTGOVOPFALIGOPVYHKMTCDSPT-----VDFECVAVHSC 197
Db 83 IPP-----RPTVYNTTTSHTSTATILNPRDYC----- 112
QY 198 FVDDGNGDTVELLNADGALDKYLLNLEYPD-----LMAQGEAHYKYKADRS- 246
Db 113 -----RGDQLHL-----LEVRDHLGRKQYGDGFLRARSSPALMAGAGKTYDFENGT 163
QY 247 -----QLFYQCOISITIT-----KEPNSRC----- 265
Db 164 LVSFTLFWEGVSLSLILIHPSGVSALMSANNOGYDRVIFMGOFVNGTSOVHSECGILT 223
QY 266 -----VRPQ-----CSEPQFGAVKTGGAAPAAQAOLRLKRSABENI-ID 309
Db 224 NNNAELCOYIDNRDQEGFYCVARPOHPCALITHMSKMKKVSYLSKQEKSLFERNSGYE 283
QY 310 VRTDINTLEISDNOALPYDLRHRALLH 338
Db 284 IMEKFWTISVSKCNLTKSVDLHESGKLDH 312

RESULT 10
```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-55

Query Match
Best Local Similarity 21.5%; Score 81.5; DB 9; Length 1948;
Matches 84; Conservative 47; Mismatches 137; Indels 123; Gaps 20;

OY 27 GEPET-ECGPTSTININTRAEGHYV-----KGLY----DQCGRNDGGHQ-- 71
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 GTPWNTATATITITWDSGNPDVSYVIEYKSKSODGPQIKEDITTTTYSIGLSPN 395
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 72 -----VAGISL-----PFDSCNVART-----RSLNPGIVTTTIVISFHLPTVK 112
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 396 SEYRIWVAVNSIGCGPSESVYRTGEQAPASPRVQARMISATMIYQMEEPVNG 455
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 113 VDRAYRVOCFYMEADKTV---SAQIEVSEITTA---FOTQIYMPVCYREIILDGPTGQ 165
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 456 LIRGFRVY-YTMEPEHFGVMNQKHNDLSLTITVGSLLDEDTYVRLAFTSVGDPLSD 514
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 166 PVQE---AIIQGVYHKWTCDSSTVTFC-----AVV--HSCFVDDGNGDTVELLNA 212
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 515 PLOVKTOGGVGPQPMNLRAARSETITLSWSPROESIIKYELLFREGDHGREV----- 569
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 213 DGCALD---KYLNNLEYPDLMAQGEAHYKYADRSQLFYQOISITIKEPNSCVRPQ 269
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 570 -GRFEDPTTSYVEDLKPNT-----YAFRLAARS----- 598
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 270 CSEPOGFAGVKTGGAAPAAQRLKKRSAPENI--IDVTDINTLEISDDNOALP 327
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 599 ---PQIGAF-----TEVVRQRTLOSQKPSAPPOVKCVSYR-----STAIL 636
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 328 VDLRHRALLHONGOPVTLAAYONGICMSPRG 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 637 VSWRPPPTTHNG-----ALVGYSVRYRPLG 662
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-738-626-5474
; Sequence 5474, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5474
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5474

Query Match
Best Local Similarity 29.8%; Score 80; DB 9; Length 293;
Matches 31; Conservative 10; Mismatches 47; Indels 16; Gaps 4;
```

```
OY 74 GILSPEDSCNVARTRSINPGIETTTTIVISFHLPTVKVDRAYRVOCFYMEADKTVSAQ 133
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 GKSLP-----AVMMTKSLQPSGVPDISDAMNIOOQOFMAVANRQ-----AAEAQKAAQ 186
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 134 IEVSEITTAFOQIYMPVCYREIILDGCPRTGQVVOFAITGPVY 177
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187 -----ATQADAPVAPQPVAMNOQFQ-APVPQGOQAPYAGAPVY 224
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-759-508B-2
; Sequence 2, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/381002
; CURRENT APPLICATION NUMBER: US/09/759, 508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 26926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-508B-2

Query Match
Best Local Similarity 18.5%; Score 80; DB 9; Length 26926;
Matches 51; Conservative 46; Mismatches 116; Indels 62; Gaps 10;

OY 107 PLEFVTVDRAVRYQCIFYMEADKTVSAQIEVSEITTAFOQIYMPVCYREIILDGPTGQ 166
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1036 PYRTTK-----PVQKLVREGSVYECQYGNKPKHYWK-----KSGVP 1075
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 167 VOFATIGOPVYHKWTCDSSTVDTPFCVAVHSCFVDDGNGDTVELLNADG-CALDKYLLNNL 225
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1076 LITGYRYKVSYNKQTECK-----LVISMTFADAGEYTIYVRNKHGFTSASASLLEEA 1129
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 226 EYPTDLMAQGEAHYKYADRSQLFYQOISITIKEPNSCVRPQCSPPGFGAVKTGGA 285
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1130 DYEL-LMKSQQ-----EMLYQVYTAFOVEPEVETAP-----GFVYSEKEY 1172
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 286 AKPAAQAQLRLKKRSAPENIIDVTDINTLEISDDNOALPVDLRH-----ALLQH 338
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1173 EKEQA-----LIRKMAKDTVAVRYVEDQEFHISFEERLIKEIYRIKTTLEELLEE 1227
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 339 NGQPVILA-----AVONGICMSPFGSMFGLSI 367
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1228 DGEKKAVIDISESEAVESGFDLRKNRYRLBGMV 1262
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-09-835-147-8
; Sequence 8, Application US/09835147
; Patent No. US20020002277A1
; GENERAL INFORMATION:
; APPLICANT: Maliszewski, Charles R.
; APPLICANT: Gayle III, Richard B.
; APPLICANT: Price, Virginia L.
; APPLICANT: Gimpel, Steven D.
; TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
; FILE REFERENCE: 2879-US
; CURRENT APPLICATION NUMBER: US/09/835, 147
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/104,585
; PRIOR FILING DATE: 1998-10-16
; PRIOR APPLICATION NUMBER: US 60/107,466
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: US 60/149,010
```

```

db 79 LFAKMEDTDEV---GWLADVLPNLISTSMGRVKNYNSLSHDT-----ATGLIQYA 128

```

QY 189 TFCVVHSCFVDDGSDTVEILNADGALD-----KYLNNLEYPPTDLMAGQEAHYV KYA 243

```

Db      79 LFAKMEYDEV-----GVALDVLSPNLISTSMLGWKNLNLMSLSDT-----ANGLOYA 128
QY      244 -----DROQFYOCQISITTKERNSECVPOQSEPOGFAVKTGGAAPK-----AAAAQLR 295
Db      129 LDQGVNATQVFEVD-----TWGMPETQARLOQSPFGIEYVTKAKADALYVVSAAISICAK 183
QY      296 LKRRSAEPENIIDVTDINTLEISDDNQALPYDLRHRALLQHNQGVIIAAVONGICMS 355
Db      184 VARDQAVKKMQFVEKIDDLDT-----DYGSGYPNDPKTKAKLKEHVEFV-----227
QY      356 PFGFSMPGSL 366
Db      228 -FGFPQFVFRFS 237

RESULT 18
US-09-881-457A-4
; Sequence 4, Application US/09881457A
; Patent No. US20020081316A1
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Cook, Stephanie M
; APPLICANT: Wild, Martha A
; TITLE OF INVENTION: No. US20020081316A1e1 Avian Herpes Virus and Uses Thereof
; FILE REFERENCE: SY01105K1OKOK
; CURRENT APPLICATION NUMBER: US/09/881,457A
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/426,352
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/804,372
; PRIOR FILING DATE: 1997-02-21
; PRIOR APPLICATION NUMBER: PCT/US95/10245
; PRIOR FILING DATE: 1995-08-09
; PRIOR APPLICATION NUMBER: 08/663,566
; PRIOR FILING DATE: 1996-06-13
; PRIOR APPLICATION NUMBER: 08/288,065
; PRIOR FILING DATE: 1994-08-09
; PRIOR APPLICATION NUMBER: PCT/US93/05681
; PRIOR FILING DATE: 1993-06-14
; PRIOR APPLICATION NUMBER: 08/023,610
; PRIOR FILING DATE: 1993-02-26
; PRIOR APPLICATION NUMBER: 07/898,087
; PRIOR FILING DATE: 1992-06-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Infectious Laryngotracheitis Virus
US-09-881-457A-4

Query Match      3.9%; Score 78.5; DB 10; Length 434;
Best Local Similarity 18.9%; Pred. No. 10;
Matches 80; Conservative 55; Mismatches 148; Indels 141; Gaps 22;

QY      45 RNAEAGHYVGLDQDECRNDGEGROVAGISLPDSCNVAKTRSLNRRGIFVTTTIVIS 104
Db      55 RNAADRHLEFNKATFTTYLLSSPASQSTAAVTYD-----ILGRRALDALITPAVG 105
QY      105 FHPLEVRKVDRAHYOCFYME-----ADKTVSAQIEVSEITTAFOQIVMPVCRYEIL 158
Db      106 PYNRLYLRVSRG-----COVELNLPISNDDMISAKKEK-----141
QY      159 DGGP-TGQVQFAIIT-----GO-----PYVHK--WTQDSEYVDTFCAY-----VHSC 197
Db      142 -GGPEASVYWFYVYIKGDDGDKYCPYRKEYRRCGDVQLSECAVQSAQMAVADYVST 200
QY      198 FVDDGNGDTVEILNADGALDKYLL-----222
Db      201 LV-SRNGAGLTFISPTAALSGOYLTLKIGRFAOTALVLEVNDRCIKISQNLFLPSKC 259
QY      223 -NNLEYPTDLMAGQEAHVYKYADRS-----QLFYQCISITIKPEPNSCVARQ 269
```

```

Db      260 WTEQYQTFG---QGEHLPLPADTNTNRHADVDYRGYEDIIQRMNLLRKKNPSADPRPD 316
QY      270 CSEPOGFAVKTGGAAPKPAAPAOQLRLKKRSAPENITVTRDINTLEISDDMQA-----325
Db      317 -SVQGEIPATYTKKAEGRTPAES-----SEKKAPF---DSEDDMQA-EASGENPALPE 366
QY      326 ---LPVDLRHRALLQ-----HNGQPVIIAAVO-----NGICMSPFGFSMGLSTALTAAY 373
Db      367 DDEVPEDTEHDDPNSPDYNDMPAVIPVEETTKSSNAVSM-PI-FAAFVACAVALVGLL 424
QY      374 IITI 377
Db      425 WMSI 428

RESULT 19
US-10-121-746-10
; Sequence 10, Application US/10121746
; Publication No. US20030036648A1
; GENERAL INFORMATION:
; APPLICANT: Miller, Andrew P.
; APPLICANT: Curran, Mark Edward
; APPLICANT: Hu, Ping
; APPLICANT: Rutger, Marc
; APPLICANT: Wang, Jian-Wang
; TITLE OF INVENTION: No. US20030036648A1e1 Human Potassium Channels
; FILE REFERENCE: SEQ-15P
; CURRENT APPLICATION NUMBER: US/10/121,746
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: US/09/336,643A
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/076,687
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/116,448
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/03826
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(646)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-121-746-10

Query Match      3.8%; Score 77.5; DB 9; Length 646;
Best Local Similarity 22.3%; Pred. No. 24;
Matches 43; Conservative 26; Mismatches 93; Indels 31; Gaps 7;

QY      199 VDDGNGDTVEILNADGCLDKYLLNLE-YPTDLMAGQEAHVYKYADRSQLEFYQCQIST- 256
Db      33 VKASRGDYVLVYVNSGRFFETW-KNTLDRYPDTLGSSSEKEFFYDADSGEYFFDRDPMF 91
QY      257 -----TIKPNNSCVRPOQSEPOGFGAVKTGGAAPKPAAPAOQLRLKKRSAPEN 306
Db      92 RHYLVNFTYTGRIKCPROECIOAFDELAFGYLV-----PELVGDCCLEYYRRKKEN 143
QY      307 IIVRTDINTLEISDDNQALPV--DLRHRALLQHNQGVIIAAVONGICMSPFGFSMPG 364
Db      144 AERLAED-EAEEQAGDGPALPAGSSLRQRLRARFENPHTSTAL-----VFYYVIGCF 195
QY      365 LSTALIAAVITTI 377
Db      196 IAVSVIANVYETI 208

RESULT 20
US-09-864-761-46862
```



```
Db 1924 TGLYISFLK-----AEPALFTSHDYKSTSHLVSRKISALAEHKVSAITLP--AE 1976
QY 273 POGFGAVKTG-----GAAKPAAAAOILRLKK-----RSAPENI 307
Db 1977 OTGWKRLKTOFNNNESYODLAVNTKDKIGVELGRLADLTLDSPILKVPILLSEPINI 2036
QY 308 IDVATDINTLEISD 321
Db 2037 ID-----ALEMRD 2044

RESULT 22
US-09-802-640-32
; Sequence 32, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsel Aruna
; APPLICANT: Kleyn Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 4563
; TYPE: PRF
; ORGANISM: Homo saplen
US-09-802-640-32

Query Match
Best local Similarity 18.7%; Score 77; DB 9; Length 4563;
Matches 81; Conservative 46; Mismatches 135; Indels 172; Gaps 20;

QY 9 CTTLLALSTSIPIVDNGVEGEPEIIBCGPTSTITINFNTRNAFEHG----- 51
Db 1662 CSDLV-----LENELNAELIGSGASMKLTITNGRFRHNAKFSLDGKALTELSTL 1710
QY 52 -----VYVKGLVD-----OEGGR--NDEGGROVAGISLPPDSCNVARTRSLNPRGI 95
Db 1711 GSAVQAMILGVDKSNITNFNEVSOEGLKLSNDMGSYA--EMKFD-----HTNSLNTAGL 1762
QY 96 FVTTTIVISEFPLVTVKDVRAVYVOCFYMEADKTVSAQISEITTAFOQIYPMPCRY 155
Db 1763 SLD-----FSSKLDNTISSDKFYKQ--TVNLQQLPYSLVYTTLNSDL-----KY 1803
QY 156 EILDDGPTG---QPVQFALIG-----QPVYHKWTCDSFTV-----DFECAY--- 193
Db 1804 NALDLTNNNGKRLLEPLKLVHAGYMLKGAYONNEIKHIAISSALASAKADTVAKQGYE 1863
QY 194 -VHSCFVD-----DGNDRIVEELNA----- 212
Db 1864 FSHRLMTDIDAGLASAIDMSNTYNSDSUHSNVRSWAPPTMTIDAHHTNGKLAIMEH 1923
QY 213 DGCALDKYLLNLEPEPTDLMAGOEAAHVYKXADRSQLFYOCQISITIKPESCEVRCQSE 272
Db 1924 TGLYISKFLTK-----AEPALFTSHDYKSTSHLVSRKISALAEHKVSAITLP--AE 1976
QY 273 POGFGAVKTG-----GAAKPAAAAOILRLKK-----RSAPENI 307
Db 1977 OTGWKRLKTOFNNNESYODLAVNTKDKIGVELGRLADLTLDSPILKVPILLSEPINI 2036
QY 308 IDVATDINTLEISD 321
Db 2037 ID-----ALEMRD 2044

RESULT 23
US-09-738-626-4782
; Sequence 4782, Application US/09738626
; Publication No. US20020197605A1
```

```
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4782
; LENGTH: 295
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4782

Query Match
Best local Similarity 20.8%; Score 76.5; DB 9; Length 295;
Matches 38; Conservative 26; Mismatches 72; Indels 47; Gaps 7;

QY 79 FDSCNNARIRSLNPRGIFVTTTVISFHL-----FVTKVDRAY--RQ 120
Db 22 FDSCE-----NPRKLKWSLTVLHDFIRDVRLTEBEWNAIDFLTRVGHITDDKQ 74
QY 121 CFYMEADKTVSAQIESEITTAFOQIIVMPYCRYEILDGP---TGQPVQFALIGQPVY 177
Db 75 EFVLVSD-FLGASMQTIVANNENAEATEDATEVTGPFVVDAPLYQNGDDIARGAVGQPRM 133
QY 178 HKWTCSEVYDFCAVHSCFVDDGNGDVEIINADGALDKYLLNLEYPPTDLMAGQA 237
Db 134 VEGT-----VKDTGPNIPNARIIEWECDEDEGLTVQYADERSAGR-A 175
QY 238 HLY 240
Db 176 HLY 178

RESULT 24
US-09-992-598-124
; Sequence 124, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavita, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
```



APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC20  
CURRENT APPLICATION NUMBER: US/09/992,598  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676

PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77; Mismatches 123; Indels 97; Gaps 19;

Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;  
QY 17 YSLPNDNGEPEETCGTSTTINFTNNAFECHYV-----KGLYDQCGRND 67  
DB 660 FSVDFDEDTSEB-LNAGGVKHL-D-STQVKMEHISTVKLSLNDTGLMEEGFKFE 717  
QY 68 GSGRVAGISLFPDSCVATRSINPRGIVTTTVISFHPLEVTKYD-----RAYR 118  
DB 718 NQRR-----NKRDRF-----FLVGNLEIRERRRLNLYPESRRCFVKVRAVR 760  
QY 119 VOCFYMEDKTVSAQIE--VSEITTAFOQIVPMPVC--RYEILDGGPTGPPVQFAIG 173  
DB 761 SERF-----LPSEQIOGVVISVILNLEPRGTGLSNPRAMGRDVSITGPNACV----- 808  
QY 174 QPVYAKMTCDSEIVTFCVAVVHSCFVDDGNGDVEIL-----NADGCAIDKYLINLEY 227  
DB 809 PAF-----CDQOSPDAVSAYVLASLA---GELQAVESSPKFNMAIGVPPYLNKNTY 859  
QY 228 -PTDLMAQEAHVYKVAARSLDFYQCOISITIKENSECVRPQCEPQG--FGAVKTCGA 284  
DB 860 RRTD--HEDRVKTKT-----FQTSMAKPRNS-----AESNGPIYAEENLRAC 902  
QY 285 AAKPAAAQLRLKRSAPENIIDVRTDINTLETISDN 323  
DB 903 EEAAPSAHFRFYQ-----ISGDRDINTVTFENED 933

RESULT 25  
US-09-989-293A-124  
Sequence 124, Application US/09989293A  
Patent No. US20020177164A1  
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Baton, Dan L.

APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijaviri, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC66  
CURRENT APPLICATION NUMBER: US/09/989,293A  
CURRENT FILING DATE: 2001-11-20  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/067770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212



Db 903 EAPPSAHPHYO-----ICGRXVNTVPFNEDD 933

RESULT 26  
US-09-989-735-124  
; Sequence 124, Application US/09989735  
; Publication No. US2002019329A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Raoul, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC61  
; CURRENT APPLICATION NUMBER: US/09/989,735  
; PRIOR FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254

;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77;  
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

OY 17 YSLPVDNGVGEPEIECGPISITINFTRNAFEHGVYV-----KGLYDQSGCRNDE 67  
DB 660 FSVDFREVTSEF- LNA GKAKVHLD-STQVKMPREHISTVKLMSLNPDTGLMEEGDPKFE 717  
OY 68 GGRVAGIASLPFSCNATRSINPRGIFVTTTVVISHFHPLEFYTKVD-----RAYR 118  
DB 718 NQRR-----NKRERT-----FLVGNLEIRERLFJNDEPSRRCCFYKVRAYR 760  
OY 119 VOCFYMEADKTVSAQIE--VSEITTAFOFOIYVMPVC--RYEILDGGPTGQVPQFAIIG 173

DB 761 SERF-----LPSEIQGVIVSINLEPRTFELSPRAMGRFDSYITPGNACV----- 808  
OY 174 QPVYHKWTCDSEIVDTFCVAVVHSCFVDDGNDTYEIL-----NADGALDKLYLNNLEY 227  
DB 809 -PAF-----CDQSPDPAVSAYVLA-----GEELQAVESSPKFPNNAIGVQPYLNNKLY 859  
OY 228 -PTDLMAGQEAHVYKRYADRSOLFVOCQISITIKEPNSFCVAPQCSPEQG--FGAVKTGCA 284  
DB 860 RRTD---HEDPRVKTA-----FQISMKAPRNS-----AESNGPIYAFENLRAC 902  
OY 285 AAKPAAQAOLRLKKRSAPENIIDVPTDITLLEISDN 323  
DB 903 EAPPSAAHFRFYQ-----IEGDRKYDNTVFPNEDD 933

RESULT 27  
US-09-990-444-124  
; Sequence 124, Application US/09990444  
; Publication No. US20020193300A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: K1javin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C19  
; CURRENT APPLICATION NUMBER: US/09/990,444  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609







[illegible]



;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77;

Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

QY 17 YSLPVDNGVEGPEICGTSITINFTNNAFGHYV-----KGLDDEGRNDE 67  
DB 660 FSYDFDEVTSEF-LNAGVYKHALD-STOVKMEPHISTYKWLMSLNDPTGLMEBEGFKKE 717  
QY 68 GGRVYAGISLPDSCNVRTSLNPRGIFVTTTVISFHLFVTKVD-----RAYR 118  
DB 718 NQR-----NKRDRT-----FLVGNLEIRERLEFLUDVPESSRCVYKRAIR 760  
QY 119 VOCFWEADKTVSAQIE---VSEITTAFOQTIVPMFVC--RREILDGPTGQPVOPALIG 173  
DB 761 SERF-----LPSEQIOGVIVISINLEPRTGFLSNPRANGRFDSVITGNACV----- 808  
QY 174 QPVYHKWDCSEVDVDFCAVHSCFVDGDCGDTVEL-----MADGCLDKTYLLNLEY 227  
DB 809 -DAF-----CDDSPDAVSYVLAISLA-----GELQAVESSPRFNPAIGVPOYLKLMY 859  
QY 228 -PTDLMAQGEAHVYKADBSQLFYQCQISITIKEPNSCEVRCQSPQSG--FGAVYTGGA 284  
DB 860 RRTD---HEDPRVYKTA-----FQISMARPRNS-----AESNSGPIYAFENLRAC 902  
QY 285 AAKPAAAOQLRLKKRSAEPENIIDVRTDINTLEISDDN 323  
DB 903 EEAAPSAHFRFYQ-----IEGDRYDYNTVPFNEED 933

RESULT 30  
US-09-991-181-124

; Sequence 124, Application US/09991181  
; Publication No. US20020197615A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC53  
; CURRENT FILING DATE: 2001-11-16  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/067770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083522  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738

```

PRIORITY APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match          3.8%; Score 76.5; DB 9; Length 1184;
Best Local Similarity 21.2%; Pred. No. 77;
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19.

QY      17 YSIPDNGVEEPEIECPTSTITINFNRMAFEGHVY-----KGLYDEGGRNDE 67
       :|::||:::||::||::||::||::||::||::||::||::||::||::||
Db      660 FSVDPERDEVTEP-LNAKKVKVHLD-STQVKMREHISTYKLMSLNPDIGLWEEBGDFKE 717

QY      68 GGROVAGISLPDSCNVARTNSLNBGRIFVTTVVISFHPLEFVTVYD-----RAYR 118
       :|::||::||::||::||::||::||::||::||::||::||::||
Db      718 NQR-----NKREDFT-----FLVGNLEIRERRLNLDPESRSRCFVKRAVR 760

QY      119 VQCFYMENDKTVSAOIE---VSEITTAFOIQIVMPVC--RREILDGGTGOPOQAIIIG 173
       :|::||::||::||::||::||::||::||::||::||::||::||
Db      761 SER-----LPSEIDQGVISVINLEPRBTGLSNBRAMGRDPDSVITGNGACV----- 808

QY      174 QPVYHKTKCTDSEVTOTFCVAVHSCFVDGNGDTVEIL-----NADGCALDKYLINLTLEY 227
       :|::||::||::||::||::||::||::||::||::||::||::||
Db      809 -PAF-----CDQOSPATYSAIYLAISLA-----GEELQAVESSPKFNNAIGAPPIYLNKLNLY 859

QY      228 -PTDLMACOEAHVYKYADRSQLFYOCQISITIKEPNSECVCPQCSEPOG--FGAVKTGA 284
       ||::||::||::||::||::||::||::||::||::||::||::||
Db      860 RRTD---HEDRPVAKTA-----FQLSMKKPRPNS-----AEESNGPIYAFENLRAC 902

QY      285 AAKPAAAOQLRLKKRSAPENIIDVRKDINTLETSDON 323
       :|::||::||::||::||::||::||::||::||::||::||::||
Db      903 EAPPSAANHFRRYO-----IESGRDYNTVPNEEDD 933

RESULT 31
US-09-993-687-124
Sequence 124, Application US/09993687
Publication No. US20020198149A1
GENERAL INFORMATION:
```

APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Eaton, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Geriltsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Kijavyn, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2730PIC11  
 CURRENT APPLICATION NUMBER: US/09/993,687  
 PRIOR FILING DATE: 2002-11-14  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77;  
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

QY 17 YSTPVNGVGEPEIECGPSTITENTRNAFESHVYV-----KGLYDEGCGRNDE 67  
DB 660 FSVDFREVTSEP-LNAGKKVHLD-STOVKMPHISTVTKLMSLNPTGLMEEGDGRFE 717  
QY 68 GGRQVAGISLPFDCSNVAFTRSLNPRGIFVTTTVVVSFHPLEFYKVD-----RAYR 118  
DB 718 NQRR-----NKREORT-----FLVGNLEIRERRLFNLDVESRRRCFYKVRAYR 760  
QY 119 VQCFYMEADKTVSAQIE--VSEITTAFOQIYPMVC--RYELIDGPGTGOVQFAIIG 173  
DB 761 SERF-----LPSQIOGVIVISINLEPRGFLSNPRAMGRFDSVITGPNAGACV----- 808  
QY 174 QPVYHKTCSEYVDFTCAYVHSCFVDGNGDIYEIL-----NADSCALDKYILANLEY 227  
DB 809 -PAF---CDQSPDAVSATVYASLA---GEELQAVESSPKRFPMNIGVQPIKLNLY 859  
QY 228 -PTDLAMGABAHYKYKADRSQLFYQCQISITIKEPNSECVRPQCSSEPG--FGAVKRTGA 284

DB 860 RRTD---HEPRVYKTA-----FOISMAKPRPNs-----AESNGPIYAFENLRAC 902  
QY 285 AAKPAAAOULRLKKSRAEPENIIDVRTDINTLEISDN 323  
DB 903 EAPPSAAHFRFYQ-----IEGDRXDVTYTFVFNEDD 933

RESULT 32  
US-09-989-734-124  
; Sequence 124, Application US/09989734  
; Publication No. US20030003531A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C64  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: US/09/989,734  
; PRIOR FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025







```

PRIORITY APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match          3.8%; Score 76.5; DB 9; Length 1184;
Best Local Similarity 21.2%; Pred. No. 77;
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19

QY 17 YSIPVDNCGVEGEPETECPTSTTINFNRNAFEGHYVY-----KGLYDQECGRNDE 67
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 660 FSVFPRDEVTSEBP-LNACKVKVYHLD-STQYVKMPREHISFVTKLWLSNLPDGLMEEGEDFKE 717
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 GGRQVAGSLPFDSCNVAFRFSLSNRGLFEYTTTVVISHPHLPVTYVD-----RAAR 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 718 NQRR-----NKKEDRT-----FLVGNLEIRERRLFNLDLPESRRCFVKVRAVR 760
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 VQCEYMEADKTVSAQIE---VSEITTAFCOTQIVPAPVC--RYEILDGPTGQPVQFAIG 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 761 SERF-----LPSQIDQGVVISVNLBERTGTGLSPRAMGRDVSITTPNGACV----- 808
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 QPVYHKTKCDESEYDTFCFVAVYHSCFVDGNGDVEIL-----NADGALDKYLLNNLEY 227
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 809 -PAF---CDQOSPDAVSAYVLASLA---GEBELAVSESPKFNNAIGVDPQVPLNKLN 859
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 -PTDLMAQGEAHVVKYKADRSOLFQYQOQISITIKEPNSCVRQCEPQG--FGAVKKTGA 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 860 RRTD---HEDRVKKTIA-----FQISMAKPRPNS-----AESNGPIYAFENLAK 902
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 AAKPAAAAQLRLKKRSAPENIIDVTDITLTLISDN 323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 903 EEARPSAAHFRFYQ-----IEGDRDYNTVPENEDD 933
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 34
US-09-997-653-124
Sequence 124, Application US/09997653
Publication No. US20030008297A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Peoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC38
CURRENT APPLICATION NUMBER: US/09/997,653
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186

```



```

; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      3.88; Score 76.5; DB 9; Length 1184;
Best Local Similarity 21.28; Pred. No. 77;
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

QY 17 YSIPVDNGVEGEPEIEGSPITSINTENTRNAFEHGVY-----KGIYDGGCRNDE 67
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 660 FSYDFRDEVSEF-LNAGKVKVHLD-STQVKMPHEHSTVKLWSLNPDTGLMEEGDKEFE 717
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 GGRVAGISLPFDCNVAFTRSINPRGIFTTTVVISFHPLEFYTKVD-----RAYR 118
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 718 NQRR-----NKRERT-----FLVGNLEIRERLFLNDVPSRRCFYKVAAYR 760
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 VQCFYHMDTVSKQLE---VSELTARFQCIYPMYVC--RYELIDGGPTGQPVQRAIIG 173
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 761 SEER-----LPBEOIQGVIVISINLEPRGFLSNPRAMGRFSVITGPGACV----- 808
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 QPYVHKWTCSEYVDFCAVHSCFVDGNGDIVEIL-----NAGSCALDXLLNLTXY 227
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 809 -PAF-----CDDQSPDAYSAIVLASLA-----GEELQAVESSPKPNPAIGVPOPYLNLNT 859
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 -PTDLNAGQEAHVYKXADRSOLFQCOISITIKEPNSECVRPOCSEFQG--FGAVYTGGA 284
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 860 RRTD---HEDPRVKTKA-----FQISMAKRPNS-----AEEENGRVAFENLRAC 902
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 AAKPAAAOURLKKRSAEPENIIDVTDINTIEISDN 323
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 903 EEAAPSAAHFRFYQ-----IEGDRYDNTVPFNEDD 933
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 35
US-09-993-667-124
; Sequence 124, Application US/09993667
; Publication No. US20030022187A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
```

```

; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC4
; CURRENT APPLICATION NUMBER: US/09/993,667
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
```



```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; PRIOR APPLICATION: 2002-04-12
; NUMBER OF SEQ ID NOS: 550
; TYPE: PRP
; LENGTH: 1184
; ORGANISM: Homo Saplen
US-10-121-049-412
```

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77;  
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

```

QY 17 YSLPVDNGVEGEPEIECGPTSTINENTRNAFEGHYV-----KGLYDGGCRNDE 67
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 660 FSYDFRDEVTSBP-LNAGKKVHLD-STQVKMPEHISTVKLMSLNDPTGLMEEGDPKFE 717
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 GGRQVAGISLPDSCNVARTRSLNPRGIFVTTTVVISFHLFVTKVD-----RAYR 118
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 718 NQRR-----NKREDRT-----FLVGNLEIRERLEMLDVPESRRCFVKYRAYR 760
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 VOCFYMEADKTVSAQIE--VSEITTAFOQTOIYPMFVC--RYEILDGPTGQPVQFAIIG 173
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 761 SERF-----LPSEQIQGVIVSYINLEPRGFLSNRANGRFDSVITGNGACV----- 808
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 QPYHHKTCDSFVDFFCVAVHSCFVDGNGDTVEIL-----NADGCLDKYLLNLEY 227
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 809 -PAF-----CDQDSPDAYSAVYLAISLA---GEELQAVESSPKFPNPAIGVPQYLKLNLY 859
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 -PTDLMAQGEAHYKKAADRSQLFYQCOISITTEKPESECVRPQCSPPQ--FGAVYKTGA 284
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 860 RRTD---HEDPRVKKTA-----FOISMAKRPNS-----AESNGPIYAFENLRAC 902
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 AAKPAAAOQLRLKRSAPENIIDVRTDINTLEISDDN 323
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 903 EEAAPSAHFRFYQ-----IEGDRYDYNIVPFNEDD 933
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

RESULT 37
US-10-123-904-412
; Sequence 412, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivartoff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; PRIOR APPLICATION: 2002-04-16
; NUMBER OF SEQ ID NOS: 550
; TYPE: PRP
; LENGTH: 1184
; ORGANISM: Homo Saplen
US-10-123-904-412
```

Query Match 3.8%; Score 76.5; DB 9; Length 1184;  
Best Local Similarity 21.2%; Pred. No. 77;  
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

```

QY 17 YSLPVDNGVEGEPEIECGPTSTINENTRNAFEGHYV-----KGLYDGGCRNDE 67
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 660 FSYDFRDEVTSBP-LNAGKKVHLD-STQVKMPEHISTVKLMSLNDPTGLMEEGDPKFE 717
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 GGRQVAGISLPDSCNVARTRSLNPRGIFVTTTVVISFHLFVTKVD-----RAYR 118
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 718 NQRR-----NKREDRT-----FLVGNLEIRERLEMLDVPESRRCFVKYRAYR 760
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 VOCFYMEADKTVSAQIE--VSEITTAFOQTOIYPMFVC--RYEILDGPTGQPVQFAIIG 173
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 761 SERF-----LPSEQIQGVIVSYINLEPRGFLSNRANGRFDSVITGNGACV----- 808
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 QPYHHKTCDSFVDFFCVAVHSCFVDGNGDTVEIL-----NADGCLDKYLLNLEY 227
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 809 -PAF-----CDQDSPDAYSAVYLAISLA---GEELQAVESSPKFPNPAIGVPQYLKLNLY 859
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 -PTDLMAQGEAHYKKAADRSQLFYQCOISITTEKPESECVRPQCSPPQ--FGAVYKTGA 284
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 860 RRTD---HEDPRVKKTA-----FOISMAKRPNS-----AESNGPIYAFENLRAC 902
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 AAKPAAAOQLRLKRSAPENIIDVRTDINTLEISDDN 323
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 903 EEAAPSAHFRFYQ-----IEGDRYDYNIVPFNEDD 933
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

RESULT 38
US-10-140-470-412
; Sequence 412, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivartoff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; PRIOR APPLICATION: 2002-05-06
; NUMBER OF SEQ ID NOS: 550
; TYPE: PRP
; LENGTH: 1184
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-412

Query Match      3.88; Score 76.5; DB 9; Length 1184;
Best Local Similarity 21.2%; Pred. No. 77;
Matches 72; Conservative 47; Mismatches 123; Indels 97; Gaps 19;

QY 17 YSLPVDNGVGEPEIEGCPISITINENTRNAFEHGVY-----KGLYDQGCGRNDE 67
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 FSYDFERDEVTEP-LNAGKKVHL-D-STGVKMPHISTVKLMSLNPTGLMEEDGRKFE 717
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 GGRVAGISLPDSCNVAFTRSNPRGIFVTTTVISFHPLEFVKVD-----RAYR 118
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 NQRR-----NKRERT-----FLVGNLEIERERLFNLDVESRCFKVKAIR 760
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 119 VQCFYMEADRTVSAQIE--VSEITTAFOQIIVPMPVC--RYEILDGGPTGQVQFAIIG 173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 SERF-----LPSEIQGVIVSVINLEPRIGFLSNPRAMGRFSDVITGPNACV----- 808
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 174 QPVIHKTCSEYVDTCAVYHSCFVDDGNGDIYEIL-----NADGALDKITLLNNLEY 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 -PAF---CDQSPDASAVYLSLA---GEEIQAVESSPKFNPNAIGVQPYLTKLNY 859
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 -PTDLMAGOBAAHYKYVADRSQLFYQCQISITIKENSECVRPQCEPQD--FGAVKKGGA 284
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 860 RRID---HEPRVKKTA-----FOISMAKPRPN-----AESNGPIYAFENLRAC 902
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 285 AAKPAAAOQLRLKKRSAPERIIDVPTDINTLEISDN 323
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 EEAAPSAAHFRFYQ-----IEGDRYDNTVTFNEDD 933
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 39
US-09-990-438-124
; Sequence 124, Application US/09990438
; Publication No. US2003002754A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Hong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertlisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC3
; CURRENT APPLICATION NUMBER: US/09/990,438
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
```

```
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
```





APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC18  
CURRENT FILING DATE: 2001-11-14  
CURRENT APPLICATION NUMBER: US/09/990,562  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/0889105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690

1	PRIOR FILING DATE: 1998-06-25
2	PRIOR APPLICATION NUMBER: 60/090694
3	PRIOR FILING DATE: 1998-06-25
4	PRIOR APPLICATION NUMBER: 60/090695
5	PRIOR FILING DATE: 1998-06-25
6	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
8	PRIOR APPLICATION NUMBER: 60/090682
9	PRIOR FILING DATE: 1998-06-26
10	PRIOR APPLICATION NUMBER: 60/090683
11	PRIOR FILING DATE: 1998-06-26
12	PRIOR APPLICATION NUMBER: 60/091360
13	PRIOR FILING DATE: 1998-07-01
14	PRIOR APPLICATION NUMBER: 60/091478
15	PRIOR FILING DATE: 1998-07-02
16	PRIOR APPLICATION NUMBER: 60/091544
17	PRIOR FILING DATE: 1998-07-01
18	PRIOR APPLICATION NUMBER: 60/091519
19	PRIOR FILING DATE: 1998-07-02
20	PRIOR APPLICATION NUMBER: 60/091626
21	PRIOR FILING DATE: 1998-07-02
22	PRIOR APPLICATION NUMBER: 60/091633
23	PRIOR FILING DATE: 1998-07-02
24	PRIOR APPLICATION NUMBER: 60/091978
25	PRIOR FILING DATE: 1998-07-07
26	PRIOR APPLICATION NUMBER: 60/091982
27	PRIOR FILING DATE: 1998-07-07
28	PRIOR APPLICATION NUMBER: 60/092182
29	PRIOR FILING DATE: 1998-07-09

Query Match	3.88;	Score	76.5;	DB	9;	Length	1184;
Best local Similarity	21.38;	Pred	NO	77;			

Matches	72;	Conservative	47;	Mismatches	123;	Indels	97;	Gaps	19;
---------	-----	--------------	-----	------------	------	--------	-----	------	-----

[illegible]

Search completed: March 24, 2003, 08:55:45  
Job time : 39 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 08:50:03 ; Search time 19 Seconds  
(without alignments)  
1958.106 Million cell updates/sec

Title: US-10-054-562a-4

Perfect score: 2026

Sequence: 1 MMIRLAFCTTIALSISIP.....ALIAVAITISFKRPNOKA 387

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*\*\*  
1: pirl:\*\*\*  
2: pirl2:\*\*\*  
3: pirl3:\*\*\*  
4: pirl4:\*\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1420.5	70.1	424	2	T20032
2	1382.5	68.1	389	2	T21239
3	1379	68.1	423	2	A49772
4	725.5	35.8	364	2	T22563
5	646	31.9	290	2	T24028
6	602.5	29.7	550	2	T23760
7	595	29.4	484	2	T20451
8	578.5	28.6	315	2	T21441
9	546.5	27.0	384	2	T2467
10	511	25.2	610	2	T22540
11	477.5	23.6	440	2	T24323
12	476	23.5	387	2	T26028
13	451	22.3	609	2	T25120
14	414	20.4	382	2	T27131
15	409.5	20.2	747	2	T34329
16	393	19.4	647	2	T23814
17	393	19.4	670	2	F88297
18	376.5	18.6	477	2	T33556
19	357.5	17.6	437	2	T34211
20	339	16.7	225	2	T33114
21	293	14.5	366	2	T25042
22	290.5	14.3	344	2	T26328
23	289.5	14.3	1286	2	A88396
24	151.5	7.5	514	2	T31736
25	141	7.0	3507	2	T34513
26	137	6.8	1262	2	T25228
27	132.5	6.5	665	2	C87789
28	120	5.9	1011	2	C87789
29	116	5.7	751	2	T21967

30	110.5	5.5	942	2	T19553	hypothetical prote
31	108	5.3	446	2	T19902	hypothetical prote
32	108	5.3	495	2	B89614	protein C43C3.3 [1
33	102	5.0	279	2	A83394	UMP-glucose-1-phos
34	101.5	5.0	828	2	T22367	hypothetical prote
35	100	4.9	935	2	S1785	peptidylglycine mo
36	94.5	4.7	875	1	URXLA2	submaxillary mucin
37	94.5	4.7	1569	2	T42233	hypothetical prote
38	93.5	4.6	649	2	T38883	carbamoyl-phosphat
39	92	4.5	2244	2	T11616	hypothetical prote
40	90	4.4	741	2	T22486	undulin 1 - human
41	90	4.4	843	2	A40970	hypothetical prote
42	90	4.4	1080	2	S48944	neural cell adhesi
43	90	4.4	1197	2	T30581	posterior-group pr
44	90	4.4	2515	2	A41519	S-layer protein pr
45	88.5	4.4	1099	2	T14850	

## ALIGNMENTS

RESULT 1  
T20032  
hypothetical protein C47G2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T20032  
R:Palmer, S.  
Submitted to the EMBL Data Library, April 1995  
A:Reference number: Z19213  
A:Accession: T20032  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1424 <M1>  
A:Cross-references: EMBL:Z49125; PIDN:CAA8934.1; GSPDB:GN00020; CESP:C47G2.1  
A:Experimental source: clone C47G2  
C:Genetics:  
A:Gene: CESP:C47G2.1  
A:Map position: 2  
A:Introns: 40/1, 115/2, 361/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein F2285.3

Query Match 70.1%; Score 1420.5; DB 2; Length 424;  
Best Local Similarity 67.1%; Pred. No. 4.8e-117;  
Matches 279; Conservative 33; Mismatches 59; Indels 45; Gaps 7;

QY	9	CTTIALSIS-IPVNGVEGEPEIECGPTSTTFNFRMAFEHVVYKGLYDQGRNDE	67
DB	8	CLAAVLASAAIPVDNVEGEPEIECGPTSTTFNFRMAFEHVVYKGLYDQGRNDE	67
QY	68	GGQVAGISLPDSCNARTRSLNPRGIEFTTYVVISFHFLEFYKVRATRVQCFYMEAD	127
DB	68	GGQVAGISLPDSCNARTRSLNPRGIEFTTYVVISFHFLEFYKVRATRVQCFYMEAD	127
QY	128	KTVSAOIESEITTAFOOTOLVMPVCRYEILDDGPGOPAFATIGOPVYHKMTCDESEY	187
DB	128	KTVSAOIESEITTAFOOTOLVMPVCRYEILDDGPGOPAFATIGOPVYHKMTCDESEY	187
QY	188	DFTCAVHSCFVDDGNDFTVEILNADSCALDKYLLNLEFPTDLMAGQEAHVYKYADRQ	247
DB	188	DFTCAVHSCFVDDGNDFTVEILNADSCALDKYLLNLEFPTDLMAGQEAHVYKYADRQ	247
QY	248	LFYQOCISITIKKEBNSCVPRQCEPQSGFAGAVKGTG-----AA	285
DB	248	LFYQOCISITIKKEBNSCVPRQCEPQSGFAGAVKGTG-----AA	285
QY	286	AKRAAA-----AOLRL-KKRS-AEPEINTDRTDINTLEISDDNQALP	327
DB	308	AAAPVAAPVAAAAAAPVAPVATLQRLRLRRKRSFGELEGILDVAVETINTLEIDEG--ASP	365
QY	328	VDLRHRLALOHNGOPVTLAAVONGICSPGFSFMGLSTALTAAVITTSFKFRP	383
DB	366	SAPEAAALVSE--ESVARRATSTGICLTPIGFASFLGIGTIVATLSTATIFYARP	419

Db 109 IGLGALROLYLNPKGLAVTTITISFHPYFTKVDRTYLNLLCRESQVYVANNISYDE 168  
 QY 139 ITTAFQOIYPMPCREIILDDGPTGQPVQAFITIGQPVYHKMTDSE-----TVDTEFC 191  
 Db 169 ISTSYVNNLMPCTGYQLISGFGFGEVEGLIGQOYVQKMDNDKMGILILKEDSFC 228  
 QY 192 AVVHSCFVDGNGDVTVEILNADGALKYLLNLEYPITDLMAGQEAHYKYADRSQLFEQ 251  
 Db 229 MMYHTCVDGDRGETSFILDSNGSIDKFLISNLEFPGNLGAGQEAHYKFRADDALEFFQ 288  
 QY 252 QOISITIKPEPNSVCVPQCEPQGFAGVKTG 282  
 Db 289 QOISITVKEPQCEVRCPEDEGGAAPVVG 319

RESULT 5  
 T24028  
 hypothetical protein R07E3.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C/Accession: T24028

R/Cottage, A.

submitted to the EMBL Data Library, May 1995

A/Reference number: Z19832

A/Accession: T24028

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-290 <MIL>

A/Cross-references: EMBL:Z49207; PIDN:CAA89068.1; GSPDB:GN00028; CESP:R07E3.3

A/Experimental source: clone R07E3

C/Genetics:

A/Gene: CESP:R07E3.3

A/Map position: X

A/Introns: 2/3; 50/2; 88/3; 127/2; 180/3; 214/1; 242/3

Query Match

Best Local Similarity 42.2%; Pred. No. 4.2e-49;

Matches 132; Conservative 50; Mismatches 79; Indels 52; Gaps 7;

QY 91 NPGIYVTVTVISFHLFVTKYDRAVRVOCFYMEADKTVSAQIEVEITTAFOQTIVPM 150  
 Db 4 SPRLFLSTNVVAFNDEFILKNDRVKVOCFYEMEMERIOKVIOISMPPTMHSKOLM 63  
 QY 151 PVRCEYELDDGPTGQPVQAFITIGQPVYHKMTDSEITDTRCAVHSCFVDGNGDVEIL 210  
 Db 64 PVRCEYELDDGPTGQPVQAFITIGQPVYHKMTDSEITDTRCAVHSCFVDGNGDVEIL 123  
 QY 211 NADGALKYLLNLEYPITDLMAGQEAHYKYADRSQLFEQOISITIKPEPNSVCVPQ 269  
 Db 124 NADGALKYLLNLEYPITDLMAGQEAHYKYADRSQLFEQOISITIKPEPNSVCVPQ 183  
 QY 270 CSEPOGFAGVKTGGAAPAAQRLILKRSAPENITDRTDINTLEISDDN--QAL 326  
 Db 184 CSEPOGFAGVKTGGAAPAAQRLILKRSAPENITDRTDINTLEISDDN--QAL 326  
 QY 327 PVDLRHRAALQH-----NGQPVIIAAYONG-----ICMSPFGSM--FMKSLIALIAAY 373  
 Db 209 HIEYEDEISDYITIPDDITISLMWLOPNDMRISLQMTAIGTTLVFLNAFLITSLV 268  
 QY 374 IITISFEKFRPNOK 386  
 Db 269 SIHVCCFRTSPK 281

RESULT 6

T23760

hypothetical protein M142.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T23760

R/McMurray, A.

submitted to the EMBL Data Library, May 1996

A/Reference number: Z19795  
 A/Accession: T23760  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: DNA  
 A/Residues: 1-550 <MIL>  
 A/Cross-references: EMBL:Z73428; PIDN:CAA97806.1; GSPDB:GN00021; CESP:M142.2  
 A/Experimental source: clone M142  
 C/Genetics:  
 A/Gene: CESP:M142.2  
 A/Map position: 3  
 A/Introns: 11/3; 39/1; 109/3; 226/1; 293/3; 341/2

Query Match 29.7%; Score 602.5; DB 2; Length 550;  
 Best Local Similarity 43.4%; Pred. No. 6.6e-45;  
 Matches 131; Conservative 48; Mismatches 108; Indels 15; Gaps 7;

QY 27 GEPEIEGPTSTINFTNRNFEHGVYVYKGLYDQGC---NDEGGRVAGISLPDSC 82  
 Db 228 GTEPICGPRIGVAKASKQPEFEGNVFVMDHYHDECRAGREKPPDSRSI-GLVVPSAC 286  
 QY 83 NVARTSLNPRGIVTTTVISFHLFVTKYDRAVRVOCFYMEADKTVSAQIEVEITTA 142  
 Db 287 NVHRTSLNPRGIVTTTVISFHLFVTKYDRAVRVOCFYMEADKTVSAQIEVEITTA 346  
 QY 143 FQOQIVPMPCREIILDDGPTGQPVQAFITIGQPVYHKMTDSEITDTRCAVHSCFV 199  
 Db 347 FQOQIVPMPCREIILDDGPTGQPVQAFITIGQPVYHKMTDSEITDTRCAVHSCFV 406  
 QY 200 DDGNGDVEILNADGALKYLLNLEYPITDLMAGQEAHYKYADRSQLFEQOISITTI 258  
 Db 407 DDGNGDVEILNADGALKYLLNLEYPITDLMAGQEAHYKYADRSQLFEQOISITTI 466  
 QY 259 K-EPNSECVRQCEPQGFAGVKTGGAAPAAQRLILKRSAPENITDRTDINT 316  
 Db 467 K-EPNSECVRQCEPQGFAGVKTGGAAPAAQRLILKRSAPENITDRTDINT 522  
 QY 317 LE 318  
 Db 523 LE 524

RESULT 7

T20451

hypothetical protein E04D5.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T20451

R/McMurray, A.

submitted to the EMBL Data Library, October 1995

A/Reference number: Z19277

A/Accession: T20451

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-484 <MIL>

A/Cross-references: EMBL:Z66496; PIDN:CAA91280.1; GSPDB:GN00020; CESP:E04D5.3

A/Experimental source: clone E04D5

C/Genetics:

A/Gene: CESP:E04D5.3

A/Map position: 2

A/Introns: 14/1; 92/3; 139/2; 253/3; 338/1; 445/1

Query Match

Best Local Similarity 29.4%; Score 595; DB 2; Length 484;  
 Matches 113; Conservative 47; Mismatches 85; Indels 6; Gaps 4;

QY 27 GEPEIEGPTSTINFTNRNFEHGVYVYKGLYDQGC---NDEGGRVAGISLPDSCVAR 86  
 Db 14 GEPEIEGPTSTINFTNRNFEHGVYVYKGLYDQGC---NDEGGRVAGISLPDSCVAR 86  
 QY 87 TRSLNPRGIVTTTVISFHLFVTKYDRAVRVOCFYMEADKTVSAQIEVEITTAFOQ 146  
 Db 74 QREIN--GVVISATVVISFHLFVTKYDRAVRVOCFYMEADKTVSAQIEVEITTAFOQ 131







[illegible]

RESULT 14  
T27131  
hypothetical protein Y53C12B.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T27131  
R:Kershaw, J.; Lennard, N.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z20316  
A:Accession: T27131  
A:Status: Preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-382 <WIL>  
A:Cross-references: EMBL:Z29278; PIDD:CA01648.1; GSPDB:GN00020; CESP:Y53C12B.6  
C:Experimental source: clone Y53C12B  
C:Genetics:  
A:Gene: CESP:Y53C12B.6  
A:Map position: 2  
A:Introns: 23/1; 87/3; 167/1; 205/3; 279/3

	Query Match	20.4%	Score 414;	DB 2;	Length 382;
	Best Local Similarity	31.6%.	Pred. No. 1.6e-28;		
	Matches 86;	Conservative	56;	Mismatches 112;	Indels 18; Gaps 7;
QY	21 VDNAGEBEIICGCTSTITINENTRNAFGHYVYGVLVDQDEGRNDEGGROVAIGSLP--	78			
Db	17 VMSDIDEGEKICATGTGITIMLETSPKGAIFLGSDAKSKCKNFSAQPSQNSIFERG	76			
QY	79 FDCSNVARTRSI-NRGIFVTVVIVISPHLEFVTRVDRAYRQCQMEADKVSAQIEFS	137			
Db	77 FDDCSRRKRQIVARCGMTMSSVLVVSVHGSIITHRDVAIQIDCYTREBNSVEFTMLSVN	136			
QY	138 EITTAFFQTQIVPMPCRYEI-LDGP-----TGOPYFAIIIGOPYHKWTCDS	185			
Db	137 APQRILIDSEKPLPCDYRVAEVETGKAAGIVTSSLSETASQAINVGDGVIHWTCSGD	196			
QY	186 T-VDFECAVNVSCFPDDNDGTVELNLANDGCALDKYLINNLEYPTDLN-AQGEAHVYRYA	243			
Db	197 APSDYCIQIVYSCEAAEDGSDTYOVAVDNGCCTTGDELLPIKYKEGSYRAAASHAFRFV	256			
QY	244 DRSOLEYCQISITIKPEPNSCVPRQSGEPQG	275			
Db	257 DNHIYFKCNIRITYKNPDGECPVNNKS-PWG	287			

RESULT 15  
 T34329  
 hypothetical protein K06A1.3 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34329  
 R:Fullon, L.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: The sequence of C. elegans cosmid K06A1.  
 A:Reference number: Z21507  
 A:Accession: T34329

```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1747 <Full>
A:Cross-references: EMBL:U03449; PIDN:AAC24299.1; GSPDB:GN00020; CESP:K06A1.3
A:Experimental source: strain Bristol N2; clone K06A1
C:Genetics:
A:Gene: CESP:K06A1.3
A:Map position: 2
A:Introns: 29/1; 62/3; 87/3; 133/3; 278/3; 430/3; 576/1; 659/3; 688/3

Query Match          20.2%  Score 409.5:  DB 2:  Length 747:
Best Local Similarity 31.4%:  Pred. No. 1e-27:
Matches    95:  Conservative    53:  Mismatches    112:  Indels    43:  Gaps    8

```

```

0Y  6 IAECTTLLNLSY-----IPVDNGVEGEIEGCGPTSTINTNTNNAEGHYVYGLYDOE 61
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  4 IDPELLMLALFNFSADPIYIEFDNHDIDDEFEVCHSGFMSLKVNTKSPSHFKVKGHRKD 63
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

0Y  62 GCRNDEGGRQVAGISLPFDSCNVANTRSNPRGIFVTTTVVISEPHLEPVTIKVBRAYRVOC 121
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  64 GC-----SESNTNANATFEESKCDVMQREANPKGNMAYATAYVQLHPLETTIKVBRAYLRIC 119
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

0Y  122 FMEADKTVSAOIE-----VSELTTFQIQIYMPYCRKEIILDGPTQ 165
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  120 FKEKEKAVGAEVSRPEFRANTNNNIIITIKCFSDTLPVQLEDESQAPCVSTYTHKESPNP 179
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

0Y  166 PVQFAIIIGOPYHHKWTCDSETVDTFGCVAVHSCFPDDENGDTEVILNADGALDKYLLNNL 225
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  180 IAKFQOLGDDVLYHVECPSE---AYQMEVYNCBDVVGGEERSKVIGENGSESDIYIMPNL 236
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

0Y  226 EXPPTD-LMAQGEAHYKKKADRSOLFYOCSI--TIKERSSECVRQC-----SE 272
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  237 IYNERTKTAFFVNSANFNFPDQNNVRISCKITVCPTL--SGTHQRPCKDASASSIDASE 293
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

0Y  273 POG 275
    | | |
Db  294 PIG 296
    | | |

```

RESULT 16  
T23814  
hypothetical protein M28.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T23814  
R:Gardner, A.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: Z19803  
A:Accession: T23814  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1 647 <ML>  
A:Cross-references: EMBL:Z49911; PIDB:CAA90131.2; GSPDB:GN00020; CESP:M28.1.1  
A:Experimental source: clone M28  
C:Genetics:  
A:Gene: CESP:M28.1  
A:Map position: 2  
A:Introns: 17/3; 37/1; 226/2; 275/3; 323/2; 453/3; 540/3; 574/1

[illegible]

```

Db 379 HQCVYKNGAGGQFEVIDQHGCTLDQLMLQPTVSEDSGMSQVDAVIFKFDPRSTVPERCT 438
QY 254 ISITFIKEPNSCEVVRPOCSEPOGFCGAVXTGGAAAKPAAAOQLRLIKKXSAEPENIIDVRTD 313
   439 ITCFSVD-DARCL-----DMPPKCGANNLRLRRKKRSHGJSSSLH 480
Db
QY 314 INTELEI-----SDNOALPVDLRHRALLQHNQPVITLAVONGICMSPEFGFSMEGELS 366
   481 ANSLTVEPDIDSKTDQMLPTPALRLKRTKEEVEYRTFCVSAFSGILISASTF--FATIS 538
Db
QY 367 IAL 369
   539 IAV 541
Db

```

```

RESULT 17
F88297
Protein M28.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F88297
R:Anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:59069613; PMID:9851916
A:Note: see websites genome.mutl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: F88297
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-670 <STO>
A:Cross-references: GB:chr_II; PIDN:CMA90131.1; PID:g38f78695; GSPDB:GN00020; CESE:M28.1
C:Genetics:
A:Gene: M28.1
A:Map position: 2

```

Query Match	19.48;	Score 393;	DB 2;	Length 670;
Best Local Similarity	29.78;	Pred. NO. 2.5e-26;		
Matches	90;	Conservative	56;	Mismatches 127;
			Indels	30;
			Gaps	6;

[illegible]

RESULT 18  
T33556  
hypothetical protein W04C9.3 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 29-Oct-1999 #sssequence\_rev1s1on 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T33556  
R:Fulton, R.; Kohlring, T.; Morris, M.  
submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid W04C9.  
A:Reference number: Z21368  
A:Accession: J33556  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-477 <FUP>  
A:Cross-references: EMBL:AF098999; PIDN:AAC68726.1; GSPDB:GN00019; CESP:W04C9.3  
A:Experimental source: strain Bristol NZ; clone W04C9  
A:Genetics:  
A:Gene: CESP:W04C9.3  
A:Map position: 1  
A:Introns: 10/3; 47/2; 140/2; 217/2; 300/1; 332/1; 458/2

RESULT 19  
T34211  
hypothetical protein F10E7.10 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34211  
R:Pauley, A.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of *C. elegans* cosmid F10E7.  
A:Reference number: 221489  
A:Accession: T34211  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-437 <PAU>  
A:Cross-references: EMBL:U41264; PTDN:AAA82426.1; CESP:F10E7.10  
C:Genetics:  
A:Gene: CESP:F10E7.10  
A:Introns: 41/3; 85/3; 266/3; 289/2; 328/2; 378/1

	Query Match	Similarity	17.6%	Score	357.5:	DB 2:	Length	437:	
	Best Local	Similarity	31.6%	Pred	N.1.8e-23:				
	Matches	87;	Conservative	47;	Mismatches	122;	Indels	19;	Gaps
Oy	1	MMIRLAFCTTLLALASYSIPVDNGVEGEPEICGPGSITININTNRNAFEGHYVVKLYDQ	60						
	:   :	:   :	:	:	:	:	:	:	:
Dd	1	MLSLVWVSFFPIYVSSLKYPNEIYDPPIVLCEDKKIKTKISMSNSNIYDGRENDD	60						
Oy	61	BCCRNDEGGROYAGISLPEDSCNVARTRSLNPRGLFVTTVTVVISPHPLFVTVKDRAYRYO	120						
Dd	61	MECVS-----RNQNKTEVAHDKCGVYNEKTEQGNGSIIRRLCLIFVQLHPLEVTESDRSYCAQ	116						
Oy	121	CFTYMADRTVSAQIEVS-EITTAFOQTQIV-----MPCKREILLDGGTGTPGVQFAITG	173						
Dd	117	CYY--ADSHVMKDISTSDISEAPLOLSPPCDAPAMPKCNSIRKOGKDGSPVOYASTIG	174						



Query Match	7.5%;	Score 151.5;	DB 2;	Length 514;
Best Local Similarity	19.0%;	Pred. No. 3.1e-05;		
Matches 67;	Conservative 63;	Mismatches 108;	Indels 115;	Gaps

hypothetical protein T23F1.5 - *Caenorhabditis elegans*  
C;species: *Caenorhabditis elegans*







C>Date: 10-May-2001 #sequence-revision 10-May-2001 #text-change 09-Nov-2001  
C/Accession: B89614  
R:anonymus, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A/Accession: B89614  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-495 <STO>  
A/Cross-references: GB:chr X; PIDN:CAA87330.1; PID:93874884; GSPDB:GN00028; CESP:C43C3.3  
A/Note: CDNA EST EMBL:T01034 comes from this gene  
C/Genetics:  
A/Gene: C43C3.3  
A/Map position: X

Query Match 5.3%; Score 108; DB 2; Length 495;  
Best Local Similarity 22.6%; Pred. No. 0.2;  
Matches 77; Conservative 47; Mismatches 128; Indels 88; Gaps 19;

Qy 31 IECGPTSTINFN-----TRNAFGHYVKGLYDOEGC---RNDEGGRQVAGISL 77  
Db 83 VCIADSTFTVVLNKSDPEVYRMISNPKSQPVYVYGHKTRHFCGSMKDEKILTNFNLT 142  
Qy 78 PRDS-CNVARTRSLNREGIFVTTVVISFH-PLFVTKVDRAVVOCFYNEADKTVSAQIE 135  
Db 143 PYGSECDVLTLDL--PKHRYAETVVLLEDNADLSFGKTRRLNHVFLCYLRANKT-----IR 196  
Qy 136 VSEITTAFC---TQIVPVCVRYEL---DGGPTGQP-----VQFIIIGQP--VYHK 179  
Db 197 FSDVSNGHVIASTGKPKP--KVEMLFRSTDSGKTLQAARENFEVFFIALSPDSAVHG 254  
Qy 180 WT-----CDSETVD-----TFCAVYHSCFVDGNGDPTVELNADGALKYLNNNE 226  
Db 255 ISPKECTSFDRDISAPPAKKITF--VGGGCFYNGMND-----IIDPLA 296  
Qy 227 YPTDLMAQGEAHVYKADRSOLFYOQOISITIKEPNSECVRPQSGFQVAKTGGA 286  
Db 297 NVNDQIVFSKFRFRFGNOSTVFHCQVYCLK--DECSK--TCYKK-----VSD 343  
Qy 287 KPAALAOGLLKKRRAEPNITDVKTINTLEISDDNAL 326  
Db 344 SNUTAERLKFRHRS-----ITDLERKRTTSPAPTDNGSL 378

## RESULT 33

A83394  
UTP-glucose-1-phosphate uridylyltransferase PA2023 [imported] - Pseudomonas aeruginosa  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence-revision 15-Sep-2000 #text-change 31-Dec-2000  
C/Accession: A83394  
R:Stover, C.K.; Pham, X.O.; Eryin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: A83394  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-279 <STO>  
A/Cross-references: GB:AE004628; GB:AE004091; NID:99948018; PIDN:AA05411.1; GSPDB:GN001  
A/Experimental source: strain PA01  
C/Genetics:  
A/Gene: galU; PA2023  
C:Superfamily: Escherichia coli utp-glucose-1-phosphate uridylyltransferase

Query Match 5.0%; Score 102; DB 2; Length 279;  
Best Local Similarity 21.9%; Pred. No. 0.3;  
Matches 49; Conservative 32; Mismatches 89; Indels 54; Gaps 11;

Qy 28 EPEIECGPTSTINFNTRNAFGHYVKGLYDOE-GCRNDEGGRQVAGISLFPDSCNVA 85  
Db 41 EEALFAGLSIEIGIVTGRKRSLEDHFDIS--YELHQIRNTDKEKYLVRRLIDECTFA 98  
Qy 86 RTTSLNPRCI---FVTTTVVISFHPFLFYTKVDRAVVOCFYNEADKTVSAQIEVSITTA 142  
Db 99 YTRQVEMKGLGNALITGRPLIDGEPVAVLADL---CLNLEGDSVLKQW---KLYNQ 151  
Qy 143 FQTOIV-----PMVCRYEILDG-----GPTGQPVQFAIIGQPVYHK 179  
Db 152 FRCSIVAIDVPEERENKKGVLAGEMIRDDIRVNTWMEKRPKEARPSMLAIIIGRIYL-- 209  
Qy 180 WTCDSFTVDFCAVYHSCFVDGNGDPTVELNADGAL 217  
Db 210 -----TPDIFDLIEQT---EPKGGEIQTIDALMKQADGCVL 244

## RESULT 34

T22367  
hypothetical protein F47G9.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 29-Oct-1999  
C/Accession: T22367; T24409  
R:White, S.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: T22367  
A/Accession: T219554  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-828 <W12>  
A/Cross-references: EMBL:Z74035; PIDN:CAA98485.1; GSPDB:GN00023; CESP:F47G9.3  
A/Experimental source: EMBL:clone F47G9  
R:Mortimore, B.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19887  
A/Accession: T24409  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-828 <W12>  
A/Cross-references: EMBL:Z74041; PIDN:CAA98523.1; GSPDB:GN00023; CESP:F47G9.3  
A/Experimental source: clone T03F7  
C/Genetics:  
A/Gene: CESP:F47G9.3  
A/Map position: 5  
A/Introns: 62/1; 181/1; 375/2; 420/1; 452/1; 585/2; 662/2; 740/3

Query Match 5.0%; Score 101.5; DB 2; Length 828;  
Best Local Similarity 20.6%; Pred. No. 1.5;  
Matches 80; Conservative 61; Mismatches 181; Indels 67; Gaps 18;

Qy 18 SIPVNGVGEPEIECGPTSTINFNTRNAFGHYVKGLYDOEGCRNDEG--GRQV 72  
Db 464 SSPLD--VKISQCTFGFISVYSPGNDNHLNVLGGEIYVRN--GHSNCSKITIGDGERI 519  
Qy 73 AGISLPEFDSNVARTRSLNREGIFVTTVVISFHPFLFYTKVDRAVVOCFYNEADKTVSA 132  
Db 520 LKIRHNDFTTC-ITKNGDIYEVVVVTVQNVESGNATVITIDQLFKVCDYNSQNAVAV 578  
Qy 133 QIEVSEITTAFC---TQIVPVCVRYEL---DGGPTGQP-----VQFIIIGQP--VYHK 179  
Db 579 AKTMRLRTTOFNKLDIYKGVNVPKPMSELRKREIT-----KARTVLTQSDIDLVTV 631  
Qy 183 DSEVYDFCAVYHSCFVDGNGDPTVELNADGALD---KYLNN--LEYPDLMAQGEA 237  
Db 632 DNST-SARHNVYKCTAVDQNGDEKILIKNGCAQNAKEVLRDEIKETATGFLPRA 690  
Qy 238 HYYKADRSOLFYOQOISITIKEPNSECVRPQSGFQVAKTGGA 286  
Db 691 FRFKGGAVKI--DEVAIVCK-----CKKVRFFVAKNFGVILFN-----N 731  
Qy 296 LKKRRAEPNITDVKTINTLEISDDNAL---PVDLRHALLQHNQPVYLAVONGI 352  
Db 732 LLDNATINIANCPNRNRTTFDESEDTTLTGNEDETKVAHLLV---QTIVDSITOLDGY 788



```

OY 242 YADRSQLFYQCQISTITKEPNSCVRPQCSFPGFV 279
Db 1442 -----EVGASFPADPNNPCISYCHN-TGEFVAV 1467

RESULT 38
T38883
hypothetical protein SPAC4H3.03c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C:Accession: T38883
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z21807
A:Accession: T38883
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-649 <MUR>
A:Cross-references: EMBL:269380; PIDN:CAA9342.1; GSPDB:GN00066; SPDB:SPAC4H3.03c
A:Experimental source: strain 972h-; cosmid c4H3
C:Genetics:
A:Gene: SPDB:SPAC4H3.03c
A:Map position: 1
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC4H3.03c

Query Match
Best Local Similarity 22.1%; Score 93.5; DB 2; Length 649;
Matches 86; Conservative 52; Mismatches 145; Indels 107; Gaps 21;

OY 25 VEDEPEEGCPITITNENRNF-----EGHYVYKGLYDOEGCRNDEGGQVAGISL 77
Db 36 LDGSVEEMCWP-----NFSPSIFARILDARAGHFSITPI-EQTSCK-----QK 78

OY 78 PFDSCNVAARTSLNPRGIEFTTVVISFH-----PLEVTKYDRA-----YRVOC 122
Db 79 YESTNIIHTRKFSERGVL---RLDFFHPRMEDYERPLMIRKYSICRIGTSRIKLECG 135

OY 123 YMEADTVSAQIEVSEIT-TAFQGTIVPMPCRYEILDGGPTQ-----PVQFAI 171
Db 136 PALDYARQSHETRYVSKITEVYQAEFVPASGDPKYLLDCVPSGQDKIDLELYPHEHL 195

OY 172 IGPVYHKKTCDBSEYVDTCAVHSCFVDGNDYVIELNADGCAIDKYLNNLEYPTDI 231
Db 196 EGGGVLSYLELEGGQITF-----TFRQEGIGLPNDVDTYTPN-----LVKLEIDSTR 242

OY 232 MAGOEAHVYKRYADRSQLFYQ---CQISITIKENPNSCV-RPQCSFPGFVAVT 281
Db 243 YNRAMTQCQVYIGRYREFYQNRNALTKLILYRPTGAVIASPTFSPLDGGVANNMDYRFT 302

OY 282 --GGAARPAQAQ-----RLKKRSAREP-NIIDVRTDINTLEISDDNQ 324
Db 303 WTRDSATFYTLAQLGFRAEAVEYMSFIYVLKKKNDGGINIV-----YSHIGDSQ 354

OY 325 AL-PVDLRH-RALLOHNGPVL--AAVON 350
Db 355 NLEVEVLTLRG--YNSHPVRIGNAVAH 382

RESULT 39
T11616
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C:Accession: T11616
R:Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z17300
A:Accession: T11616
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2244 <BAR>
A:Cross-references: EMBL:254328; NID:g1009451; PID:g1009456

```

```

A:Experimental source: strain 972h(-)
C:Genetics:
A:Map position: IL
A:Note: SPAC22G7.06c
C:Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology;
arabamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-p
F:61-1519/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F:61-440/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain hom
F:265-440/Domain: trpG homology <TRG>
F:476-1516/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain h
F:476-926/Domain: biotin carboxylase homology <BCB>
F:1532-1857/Domain: Bacillus dihydroorotase homology <DHO>
F:1937-2237/Domain: aspartate/ornithine carbamoyltransferase homology <AOC>

Query Match
Best Local Similarity 20.3%; Score 92; DB 2; Length 2244;
Matches 74; Conservative 58; Mismatches 153; Indels 80; Gaps 17;

OY 42 FTRNAFEGHVYKGLYDOEGCRNDEG-----GROVAGISLPFDSCNVAARTSLNPRGI 95
Db 999 YTYNVAEHDHIF-----NDKGVWVIGSVYRIGSSVYFDMCAVRAVETLDRGV 1048

OY 96 FVTTVVISFHPLEVT---KYDRAVYQCFYWEADKTVSAQIEVSEITTAFOGT---V 148
Db 1049 ---KTIIMVYNPETVSTDYDEADRLY-FENIGLETVLIDYEGSSGIIIMGGGTANNI 1104

OY 149 RMPVYCV-----EIIDGGT---GQPVFAIIGQPVYHKKTCDBSEYVDTCAVVA 195
Db 1105 ALPLHRENVKILGTSPMEIDGAEHRRKFSRMLDIDGVDPKKELTSFDE-ADKFC--- 1159

OY 196 SCFVDGNGDTV--EILNADGCAIDKYLNNLEYPTDI-----MAGOEAH--YKYA 243
Db 1160 -----DYGYPLVLRPSYVLSGAAMTYVSGDLSYQQAIAIKDKRPVYSKI 1210

OY 244 DNSQLFYQCQISTITKEPNSCVRPQCSFPGFVAVTGA--AKPAAQAQLRLKK 299
Db 1211 ENAK-----EIEDAVAREGKVMHVISHEVENAGVSHGADATLVLPOLAPFTTERIVD 1265

OY 300 RSEAPENIIDVRTDINTLEISDDNQALRPDLRHALLOHNGPVLIAAONICSPFG 359
Db 1266 AAKAGELNITGPYVITQFIARDNEIKYECNVRA---SRSPFVSKYL--GVDMISMAT 1320

OY 360 SMFMG 364
Db 1321 DVIMG 1325

RESULT 40
T22486
hypothetical protein F52B11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22486
R:Matthews, L.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19569
A:Accession: T22486
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-741 <WIL>
A:Cross-references: EMBL:282268; PIDN:CAB05199.1; GSPDB:GN00022; CESP:F52B11.3
A:Experimental source: clone F52B11
C:Genetics:
A:Gene: CESP:F52B11.3
A:Map position: 4
A:Introns: 48/L; 298/L; 341/3; 690/2

Query Match
Best Local Similarity 22.1%; Score 90; DB 2; Length 741;
Matches 89; Conservative 43; Mismatches 181; Indels 90; Gaps 20;

OY 22 DNGVGEPEIE-----CGPTSIITNENRNFEGHVYKGLYDOEGCRN- 65

```

```
Db 366 DANIKANPSMDYEFNCESGFGMALCTNEGIRFIVNTKEPYTGAIYAAERFST--CSQV 423
QY 66 DEGRQVAGISLPEDSCNVATRSLNPRGIVTTTVISFHPLE---VTKVDRAVYQC 121
Db 424 VENAKQ--ISTPPPTVSSDCGTIRDGKMEALVWVSLDGLPHOVTTMWDPRFVYSC 480
QY 122 -----FYMEADKTVSAQIEVSEITTFAFQTIIVMPV-CRYEILDGCPGQPVQFAITGQ 174
Db 481 DVSMDKKVKKEGSVYVTTIYEASSONTVYLDVATPPVSAELQIILN--QLEEPLIKASIGD 538
QY 175 PVYHKWTCDESETVDFCAVHSCFVD--DGNQDVEI-LNADGCALDKYLLNMLEYPTDL 231
Db 539 PLLIVIT--SEQAGPHNMWVTECTATRVGFGDVTVPFTLIENGCP-----RYPA-L 586
QY 232 MAGOEAHVYKADRSQI-----FYOCQISITIKERNSECVRPOCSEPOGFGAVKTGGA 284
Db 587 VGPVEQDFDKNRILKSDLRARFLDGSYDQIVCSIM-----FCAGPNQCPVSNCLDS 637
QY 285 AAKPAAQAQLRLKKRSAEPENIIDVPTDINTLEISDDNQALPYDLR-----HRALL 336
Db 638 GTNELFMHGR--KKRSAD-----LNGETEEKLSAIIIRVFAKGEDEEEMEM 682
QY 337 OHNGQPVILAAVONGICMS--PFGFSMFUGLSIALIA-AVITTI 377
Db 683 ANNTMTMSDSTELICIAEPFVSVVSLVLCFALSAITAI 725
```

Search completed: March 24, 2003, 08:51:05  
Job time : 30 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 08:49:12 ; Search time 13 Seconds

(without alignments)  
1234.718 million cell updates/sec

Title: US-10-054-562a-4

Perfect score: 2026

Sequence: 1 MMIRLAFCTTIALSYSIP.....ALIAAVITRISFKFRNQKA 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1420.5	70.1	424	1	CUTL_CAEEL
2	108	5.3	446	1	YOH3_CAEEL
3	102	5.0	279	1	GAU1_PSEAE
4	96	4.7	279	1	GAU2_PSEAE
5	94.5	4.7	875	1	AMD_XENLA
6	93.5	4.6	649	1	YAY3_SCHPO
7	92	4.5	2244	1	PYR1_SCHPO
8	90	4.4	1080	1	NRK1_YEAST
9	90	4.4	2515	1	TUD_DROME
10	88	4.3	429	1	THRC_BUCAT
11	86.5	4.3	407	1	CGEL_CHICK
12	86.5	4.3	1260	1	YAOE_SCHPO
13	86	4.2	404	1	ISCS_METTE
14	85.5	4.2	1117	1	DSG2_HUMAN
15	85	4.2	1025	1	HIRA_FUGRU
16	85	4.2	4485	1	DYHG_CHIRE
17	84.5	4.2	410	1	CAOC_CANFA
18	84	4.1	2193	1	POL6_EC12T
19	83.5	4.1	563	1	MUC6_BOVIN
20	83	4.1	666	1	TKT_BACHD
21	83	4.1	820	1	PDZ1_ARATH
22	83	4.1	1627	1	PAPA_HUMAN
23	81.5	4.0	430	1	SYN_BACHD
24	81.5	4.0	2286	1	DPOE_HUMAN
25	81	4.0	848	1	TORA_ECOLI
26	81	4.0	1080	1	CARB_XYLA
27	81	4.0	1663	1	CO3_MOUSE
28	80.5	4.0	311	1	RLAO_CAEEL
29	80.5	4.0	430	1	SYN_STAM
30	80.5	4.0	864	1	LOX2_PEA
31	80	3.9	848	1	TORA_ECO57
32	80	3.9	974	1	AMD_HUMAN
33	79.5	3.9	340	1	PURA_METTH

34	79.5	3.9	1753	1	YEW2_YEAST	P32634 saccharomyc
35	79.5	3.9	2214	1	PYR1_YEAST	P07259 saccharomyc
36	79	3.9	553	1	YF61_METJA	O58956 methanococc
37	79	3.9	750	1	CTPB_MYCLE	P46840 mycobacteri
38	79	3.9	827	1	CADH_MOUSE	O91100 mus muscula
39	79	3.9	3587	1	TYCB_BACBR	O30408 b tyrocidin
40	78.5	3.9	299	1	RNHL_HUMAN	O75792 homo sapien
41	78.5	3.9	436	1	RP54_BACSV	P24219 bacillus su
42	78.5	3.9	556	1	GLI_CHICK	P55878 gallus gall
43	78.5	3.9	570	1	ZPR_FELCA	P48834 felis silve
44	78	3.8	507	1	RREP_MEASI	P26033 measles vir
45	78	3.8	785	1	SM3E_CHICK	O42237 gallus gall

## ALIGNMENTS

RESULT 1	ID	CUTL_CAEEL	STANDARD	PRT	424 AA.
AC	003755	O18693:			
DT	01-FEB-1994	(Rel. 28, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	15-DEC-1998	(Rel. 37, Last annotation update)			
DE	Cuticlin 1 precursor.				
GN	CUT-1 OR CATG2.1.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;				
OC	Rhabditidae; Pelodierinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-Bristol N2;				
RX	MEDLINE=91323673; PubMed=1864469;				
RA	Sebastiano M., Lassarano F., Bazzicalupo P.;				
RT	"cut-1 a Caenorhabditis elegans gene coding for a dauer-specific				
RT	noncollagenous component of the cuticle.";				
RL	Dev. Biol. 146:519-530(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-Bristol N2;				
RA	Palmer S.;				
RL	Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: COMPONENT OF THE CUTICLES. IT FORMS A RIBBON				
CC	APPROXIMATELY 2 MICRONS WIDE RUNNING ALONG THE LATERAL LINES				
CC	UNDERNEATH THE ALAE. CONTRIBUTES TO THE FORMATION OF EXTRACELLULAR				
CC	ENVVELOPES PROTECTING THE ORGANISM FROM THE ENVIRONMENT.				
CC	-1- SUBCELLULAR LOCATION: LOCALIZED IN A NARROW BAND NOT EXPOSED				
CC	TO THE EXTERIOR. RUNNING UNDERNEATH THE ALAE.				
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED DURING DAUER LARVA FORMATION.				
CC	-1- DOMAIN: THE SMALL REPEATS A-A-P-[AVI] ARE ALSO PRESENT IN MANY				
CC	PROTEINS CONSTITUTING THE PROTECTIVE ENVELOPE OF OTHER SPECIES.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL; M55997; AAA27995.1; ALT_INIT.				
DR	EMBL; Z49125; CAA86934.1; -.				
DR	PIR; A49772; A49772.				
DR	Wormpep; C4762.1; CE02165.				
DR	Interpro; IPR001507; Endoglin/CD105.				
DR	SMART; SM00241; Zp. 1.				
KW	Cuticle; Signal; Repeat.				
FT	SIGNAL	1	18		POTENTIAL.
FT	CHAIN	19	424		CUTICLIN 1.
FT	DOMAIN	302	323		4 X 4 AA REPEAT OF A-A-P-[AVI].
FT	REPEAT	302	305		1.

```

FT REPEAT 307 311 2.
FT REPEAT 312 315 3.
FT REPEAT 320 323 4.
FT CONFLICT 27 27 G -> P (IN REF. 1).
FT CONFLICT 29 29 P -> G (IN REF. 1).
FT CONFLICT 306 306 A -> G (IN REF. 1).
FT CONFLICT 309 309 MISSING (IN REF. 1).
FT CONFLICT 389 390 CL -> SS (IN REF. 1).
FT CONFLICT 395 396 FA -> LP (IN REF. 1).
FT CONFLICT 401 402 IG -> MR (IN REF. 1).
SQ SEQUENCE 424 AA; 45168 MW; 2D178A602C1500B5 CRC64;

```

Query Match  
 Best Local Similarity 70.1%; Score 1420.5; DB 1; Length 424;  
 Matches 279; Conservative 33; Mismatches 59; Indels 45; Gaps 7;

```

QY 9 CTTLLALSYS-IPVDNGVEGEPEIEGCPSTITNFNTRNAFEGRVYVKGLYDDEGRNDE 67
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8 CLAAVLASAIIPVDNNVESEPEVEGCPNSTVYNTNRPEBGHVYVKGLYDQAGCRSDE 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 GGRQVAGISLPDSCVAVARTSLNPGIYVTTTVISFHPLEFYTKVDRAVRYOCFYMEAD 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 GGRQVAGIELPFDSCVARTSLNPGIYVTTTVISFHPLEFYTKVDRAVRYOCFYMESD 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 128 KTVSAQIEVSEITAFOTQIVPMPVRCYREITLGGPTGQVQAFALIGQPYVHKTKDSEVY 187
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 128 KTVSTQIEVSDLTITTAQTOVVPMPVCKYELLDGSPSGOPTQVATIGQVYHKTKDSEVY 187
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 188 DFFCAVYHSCFVDGNGDTVEILNADGALKYLLNNLEPYTDLMGOGAHYKYKADRSG 247
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 188 DFFCAVYHSCFVDGNGDTVEILNADGALKYLLNNLEPYTDLMGOGAHYKYKADRSG 247
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 248 LFYQCOISTITKEPNSECVRCPCSEPOGFGAVKATGCG-----AA 285
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 248 LFYQCOISTITKEPNSECVRCPCSEPOGFGAVKATGCG-----AA 285
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 286 AKPAA-----AQLRL-KKRS-AEPENIDVPTDINTLEISDNGALP 327
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 308 AAPVAPVAAAAAARVPATLAQLRLKKRSFEGNEILDVRYVINTLDMEG--ASP 365
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 328 VDLRRALLQHNGQPYITLAAYONGICMSPFGSMGSLIALIAVYITISFKFRP 383
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 366 SAPEAAALVSE--ESVRRRATSTGICLPFGASFICGTYVATLMSATIFVVARP 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 2  
 YOH3-CAEEL STANDARD; PRT; 446 AA.  
 AC 009276;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 50.0 kDa protein C43c3.3 in chromosome X.  
 GN C43c3.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderiinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Sulston J.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RM [2]  
 RP REVISIONS.  
 RA Durbin R.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL: 247067; CAAB7330.2; -
DR WormRep: C43C3.3; CE23591.
DR InterPro: IPR001507; Endoglin/CD105.
DR SMART: SM00241; ZP; 1.
KW Hypothetical protein; Transmembrane; ATP-binding.
FT TRANSMEM 377 397 POTENTIAL.
FT NP_BIND 122 129 ATP (POTENTIAL).
SQ SEQUENCE 446 AA; 49954 MW; ACC4AB69B30032D84 CRC64;

```

Query Match  
 Best Local Similarity 5.3%; Score 108; DB 1; Length 446;  
 Matches 77; Conservative 47; Mismatches 128; Indels 88; Gaps 19;

```

QY 31 IEQGPRSTINRN-----TRAFSGHYVYVGLYQDEG---RNDEGQVAGISL 77
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 34 VDCIADSFVVLNKSDEPMYRMISNPKSQPVYVYGHKTRHPCGTSMKDEKLTJNFNLT 93
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 78 PFDG-CNVARTSLNPGIYVTTTVISFHP-PLFYTKVDRAVRYOCFYMEADKTVSAQIE 135
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 94 PYGSECDVTLTDL-PKHRYAETTVLEDMADLSGKTRLMNHVCLYRANKT-----IR 147
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 136 VSEITTAFO-----TQIVPMPVCRTEIL---DGCPGCP-----VQFAITQP--VYHK 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 148 FSDVSNQHEVIASTGKPKP--KVEMLFIRSTSGKTLQAARENEVEFFIALSPDSAYHG 205
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 180 WT-----CDSEFVD-----TFCAVYHSCFVDGNGDTVEILNADGALKYLLNNLE 226
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 206 ISPKECTFSDREDISAPDAKKTTF--VQGGCFVNGKND-----ITDPLA 247
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 227 YPTDLMAQGAHAYKYADRSQLEFYQCOISTITKEPNSECVRCPCSEPOGFGAVKATGAAA 286
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 248 NVNDQYIFSKFRTFRGNGSTVFVHCQVQVCLK--DECSK-TCYKK-----VSD 294
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 287 KRAAAQLRLKKRSAPENIIDVPTDINTLEISDNGAL 326
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 295 SMLTAERLFRHKRS---ITDLERRTTRSAPTDNGSL 329
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 3  
 GAUL\_PSEAE STANDARD; PRT; 279 AA.  
 ID GAUL\_PSEAE  
 AC 091291;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) (UDP-glucose  
 pyrophosphorylase) (UDPP) (Alpha-D-glucosyl-1-phosphate  
 uridylyltransferase) (uridine diphosphoglucose pyrophosphorylase).  
 GN GAUL OR PA2023.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PAOI;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an  
 opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN STATIONARY PHASE SURVIVAL (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: UTP + alpha-D-glucose 1-phosphate =  
 CC diphosphate + UDP-glucose.





```
DR HSP: P14925; 1PHM.
DR InterPro: IPR000333; Cu2_monooxygenase.
DR InterPro: IPR001258; NHL.
DR InterPro: IPR00720; Pamonoxigenase.
DR Pfam: PF01082; Cu2_monooxygen; 1.
DR Pfam: PF01436; NHL; 4.
DR Pfam: PF03712; Cu2_monoox_C; 1.
DR PRINTS: PR00790; PAMONOXGNASE.
DR PROSITE: PS00084; CU2_MONOOXYGENASE_1; 1.
DR PROSITE: PS00085; CU2_MONOOXYGENASE_2; 1.
DR Olfactorin; Monooxygenase; Copper; Vitamin C; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 39
FT CHAIN 1 875
FT DOMAIN 40 763 POTENTIAL.
FT TRANSMEM 764 787 MONOOXYGENASE II.
FT DOMAIN 788 875 INTRAGANULAR (POTENTIAL).
FT DOMAIN 358 364 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 387 390 POLY-MET.
FT DOMAIN 852 856 POLY-GLU.
FT CAROHD 455 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 743 743 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 875 AA; 97084 MW; C07373AF6BF13450 CRC64;

Query Match
Best Local Similarity 4.7%; Score 94.5; DB 1; Length 875;
Matches 50; Conservative 25; Mismatches 66; Indels 91; Gaps 12;

OY 33 CQPTSTINFTNFAFEHGYVKGLYDQGCGRNDEGQVAGISLPFDSQNVARTSLNP 92
DB 530 CQPTVAADPTGTFNFDVADGY-----CN-SRIMQFSP 560
OY 93 RGEIV-----TTTVVISFHPFYTKVDRARYVQCFYMEADKYVS 131
DB 561 NCMFTIMQGEETSSNLPFGCFRIPHSLSMISDQGLCVADENG-RICGFIAKGEFVK 619
OY 132 AOIEVSEI-TTAFQTOIVPMPCREIILDGPTGQVQFAIGQPVYHKMTCDSETVDTF 190
DB 620 -QIKHQEFGREVFAYSYAP-----GG-----VLAVNGKPYG---DSRPVQGF 659
OY 191 CAVVHSCFVDGNGDPVELINADGALDKYL--LNINLEPPTLMAGQEHVY 240
DB 660 -----MLNSNGD-----ILDTFIPARKNFEMPHDIAAGDQTVY 694

RESULT 6
YAY3_SCHPO STANDARD: PRT; 649 AA.
AC Q10211;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CAH3.03c in chromosome I.
CN SPAC4H3.03c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
```

```
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filicz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurt S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RA Nature 415:871-880(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 269380; CAA93342.1; -.
DR KM Hypothetical protein; Transmembrane.
FT TRANSMEM 265 285 POTENTIAL.
FT TRANSMEM 564 584 POTENTIAL.
FT TRANSMEM 626 646 POTENTIAL.
SQ SEQUENCE 649 AA; 74488 MW; 0C97C10E603EB3D CRC64;
```

```
Query Match
Best Local Similarity 4.6%; Score 93.5; DB 1; Length 649;
Matches 86; Conservative 52; Mismatches 145; Indels 107; Gaps 21;
```

```
OY 25 VEGPEIEEGPSTITINFTNFAFEHGYVKGLYDQGCGRNDEGQVAGISL 77
DB 36 LDGSVEAMCWP-----NDSPSIFARILDARAGHSITPI-EQTSCK-----QM 78
OY 78 PFDSCNVARTRSLNPGIFVTTVVISFH-----PLFVTKVDRA-----YRVOCF 122
DB 79 YEPSTNIIHTKPYSEKGLV---RLIDFFHPEWDEPEPLPMIRVSGISRGSRKILKCF 135
OY 123 YMEADKYVSAQIEVSEI-TTAFQTOIVPMPCREIILDGPTGQ-----PVQFAI 171
DB 136 PALDYARQSHETRYKSKITENYQAEFVPSAGDPKYYILDCVPSDQILKIDELIYPAEHLI 195
OY 172 IGQPVYHKMTDSEVDVDFCAVHSCFVDGNGDPVELINADGALDKYL--LNINLEPPTLM 231
DB 196 EGGGYISLLEBEGOEITF-----TFQEGGSGPVDVDTN-----LVDKLEDSYKR 242
OY 232 MAGQEAHYKYADRSQLFYQ---COISITIKENPSECV-RPQCSPEQGFCAVKT----- 281
DB 243 YWRAMIQCCVYTGRRREVQNALTKLIIETPGAVIASPTFSLPDLGGVGRMWDYFT 302
OY 282 --GGAAPAAQAOL-----RLIKRSAPPE-NIIDVRTDITLEISDNO 324
DB 303 WIRDSAFPIYALQDLGFRAEAVVEYMSFTYHVLKRNKDGGINIV-----YSIHGDSQ 354
OY 325 AL-PVDLRH-RALLQHNQAPVIL--AAVON 350
DB 355 NLEEVELTHNG--YNSHPVIRIGMAVHR 382

RESULT 7
PYR1_SCHPO STANDARD: PRT; 2244 AA.
AC Q09794;
DT 01-NOV-1995 (Rel. 32, Created)
```



DB 1160 -----DIYGYPLVLRPSVYLSGAAMTYISQSDLSYIQQAVAIKQHPVITSKIY 1210  
 QY 244 DRSQLFYQCOISITIKKEPNSECVRPQCSPPGFGAVKGTGAA---AKPAAAOQLRLKK 299  
 DB 1211 ENAK-----EIELDVAAREGKVMHYISEHVENAGVSGDATLVLPQDLAPTIERIYD 1265  
 QY 300 RAEEPENITDVATDINTLEISDNOALPYDLRHRALLQHNQGVPIIAAVQNGICSPFEF 359  
 DB 1266 AAKKGALGNITGPYNIQPIARDNEIKYIECNVRA---SRSEPFVSKVI--GVDMTSMAT 1320  
 QY 360 SMFMG 364  
 DB 1321 DYIMG 1325  
 RESULT 8  
 NRK1\_YEAST  
 ID NRK1\_YEAST STANDARD: PRT: 1080 AA.  
 AC P38692: 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase NRK1 (EC 2.7.1.37) (N-riich kinase 1).  
 GN NRK1 OR K1C1 OR YHR102W  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=DC-5;  
 RA Fukami Y.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 Du Z., Favell A., Fulton L., Gattung S., Giesel C., Kisten J.,  
 Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,  
 Lacroix P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,  
 Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaidin M.;  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 VIII";  
 RL Science 265:2077-2082(1994).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- SIMILARITY: BELONGS TO THE SRR/THR FAMILY OF PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D29980; BA06250.1; -;  
 DR EMBL; U00059; AAB68860.1; -;  
 DR PIR; S48944; S48944.  
 DR SGD; S0001144; K1C1.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_Pkinase.  
 DR Pfam; PF00069; Pkinase.1.  
 DR ProDom; PD000001; Euk\_Pkinase.1.  
 DR SMART; SM00220; S\_TKc.1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST.1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM.1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 23 276 PROTEIN KINASE.  
 FT NP\_BIND 29 37 ATP (BY SIMILARITY).

FT BINDING 52 52 ATP (BY SIMILARITY).  
 FT ACT\_SITE 144 144 BY SIMILARITY.  
 SQ SEQUENCE 1080 AA; 117061 MW; 9989EAF315E0B94 CRC64;  
 Query Match 4.4%; Score 90; DB 1; Length 1080;  
 Best Local Similarity 21.5%; Pred. No. 7.4;  
 Matches 45; Conservative 41; Mismatches 93; Indels 30; Gaps 6;  
 QY 144 QYQIYPMPCVRCIILDGCP-----TGQPVQPAITIGQPYHYHMTGDSFEVDFCAVHSC 197  
 DB 200 KYDWSLIGITTYEIAITGNPNPCDVEALRMQLIKSPPRLEDRSYSTLKKEFIAL---C 256  
 QY 198 EYDQNGDVTVEILNADGCAIDKYLLNNLEYPTDLAAGQEAHYKXADRSQFYQCOISIT 257  
 DB 257 IDEDPK---ERLSADDLKSKFIRAHKATPTSILKELISRYLLFRDKNKKNKKEIGSIP 312  
 QY 258 IKEPNSCEVRPQCSPPGFGAVKGTGAAAKPAAAOQLRLKRSAPENIIDVRIDINTL 317  
 DB 313 ENEPSKPEAPKPSQ-----NGGDEAKQKSIASNDNEIKRVN---EGDVEMKWDPSL 362  
 QY 318 EISD-----DNOALPYDLRHRALLQHN 339  
 DB 363 SSSDYIENINIMDLALDENNEMATAQHD 391  
 RESULT 9  
 TUD\_DROME  
 ID TUD\_DROME STANDARD: PRT: 2515 AA.  
 AC P25823:  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Maternal tudor protein.  
 GN TUD.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92038995; PubMed=1936993;  
 RA Golubetski G.S., Bardsley A., Tax F., Boswell R.E.;  
 RT "Tudor, a posterior-group gene of Drosophila melanogaster, encodes a  
 RT novel protein and an mRNA localized during mid-oogenesis";  
 RL Genes Dev. 5:2060-2070(1991).  
 CC -1- FUNCTION: REQUIRED DURING OOGENESIS FOR THE FORMATION OF  
 CC PRIMORDIAL GERM CELLS AND FOR NORMAL ABDOMINAL SEGMENTATION.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE LIFE CYCLE.  
 CC -1- MISCELLANEOUS: THE TUD MRNA ACCUMULATES DURING THE EARLY TO MIDDLE STAGES  
 CC OF OOOGENESIS.  
 CC -1- SIMILARITY: CONTAINS 9 TUDOR DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X62420; CAA44286.1; -;  
 DR PIR; A41519; A41519.  
 DR HSSP; O16637; IG5V.  
 DR FlyBase; FBgn00038901; tud.  
 DR InterPro; IPR001097; Maternal\_tudor.  
 DR InterPro; IPR002999; Tudor.  
 DR Pfam; PF00567; TUDOR.10.  
 DR SMART; SM00333; TUDOR.9.  
 DR PROSITE; PS50304; TUDOR.9.  
 KW Developmental protein; Repeat.  
 FT DOMAIN 455 513 TUDOR 1.

```
FT DOMAIN 641 696 TUDOR 2.
FT DOMAIN 1062 1122 TUDOR 3.
FT DOMAIN 1355 1414 TUDOR 4.
FT DOMAIN 1662 1718 TUDOR 5.
FT DOMAIN 1839 1898 TUDOR 6.
FT DOMAIN 2023 2082 TUDOR 7.
FT DOMAIN 2211 2269 TUDOR 8.
FT DOMAIN 2392 2451 TUDOR 9.
SQ SEQUENCE 2515 AA; 285236 MW; 683C100AD308BADA CRC64;

Query Match 4.4%; Score 90; DB 1; Length 2515;
Best Local Similarity 19.1%; Pred. No. 23;
Matches 70; Conservative 51; Mismatches 140; Indels 106; Gaps 16;

QY 16 SYSTIPVNGVGEPEIEICGPTST---INFNTRNAEG--HYVYKGLYDOEGCRNDEGG 69
DB 1904 SLOAP-DAVISWPEAKRAELTGEGLVFTQLKPGDHYVITDILLD-----G 1953
QY 70 ROVAGISLPDSCNVAFTSRINPRGIFVTPTTVISFHPLEFTKYVDRAVYOCFYMEAD-- 127
DB 1954 ENITDRLPL--CORKEPKRSKESLAVTTKAIT-----THVENTSRITLQFSEKDSL 2004
QY 128 -----KTVSAQIE-----VSEITTAFOQIIVMPYCRY 155
DB 2005 MDICEKLNKSGKIQPKTEKAAVDMCVQFADLEFYRSRLLEVLEDDQYKVI----- 2057
QY 156 EILDGCTP-----GQPVQFAITIGQPVYHKWTCDSF-----TYDTFCVAVVHSC 197
DB 2058 -LDYGMTTVVDKLYELPQEFLLI-KPYAE--ICSMPSAIFERNKALITLTPFALDSC 2113
QY 198 FVDDGNDGTYEILNADGALDKYLLNNLEPYPTDLMAGOEAHVYKYVARSOLFQCOISIT 257
DB 2114 -----KGVAVAEFVNKSASPPVYVRLTKDKRSLKLTIEHLQKLYVAELKLI 2158
QY 258 IK-EPNSECVRPQCEPQGFQAVKGTGAAPAAQRLILKRSAPENIIDVTDINT 316
DB 2159 QKRENSECEITISYNSPKSFYQMKHNSADULLIVTKLQSLKREKL--KKLIDPTNSNG 2216
QY 317 LEISDDN 323
DB 2217 VCYSQED 2223

RESULT 10
THRC_BUCAI STANDARD; PRT; 429 AA.
AC P57289;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Threonine synthase (EC 4.2.3.1).
GN THRC OR BU192.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OC Bacterial; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. Aps.";
RL Nature 407:81-86(2000).
CC -1- CATALYTIC ACTIVITY: O-phospho-L-homoserine + H(2)O = L-threonine +
CC phosphate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: Threonine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE SERINE/THREONINE DEHYDRATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001118; BAB12909.1; -.
DR InterPro: IPR001926; B6_enzyme_beta.
DR InterPro: IPR000634; S/T_dehydrase.
DR InterPro: IPR004450; Thr_synthase.
DR Pfam; PF00291; PALP; 1.
DR TIGRfams; tigr00260; thrc; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Threonine biosynthesis; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 108 108 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 429 AA; 48659 MW; F66C73BEF5FAFC6 CRC64;

Query Match 4.3%; Score 88; DB 1; Length 429;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 78; Conservative 57; Mismatches 145; Indels 110; Gaps 17;

QY 14 ALSTIPVNGVGEPEIEIC-----GPTSTINFNTR----- 45
DB 79 AFSEKHPLK--IKITKDIHCELEFHPGLAFKDGARFMAQMLLNKKNESVTILRATS 136
QY 46 -----NAFEGH-----VYVKG-----LYDGGCRNDEGGROVAGISL--PFDSCNV 84
DB 137 GDTGAANAANAYGKKNRVITLLYPRKISLQELFCTL--GNIKITISNGSEDDCK 193
QY 85 ARTSLNPRGIFVTTVVISFHPLEVTKYVDRAVYOCFYMEADTVSAQIEVSEITTAQ 144
DB 194 LVKEAFNDKILKES---IGLSANSINISRLAQICYFFAFSLISEQR-----K 241
QY 145 TQIVPMPCREIILDGCTGQPVQFAITIGQPVYHKWTCDS--SEYVDTFCVAVVHSCFVDDG 202
DB 242 NLVIAVPCNGNLTAGLSKS---LGLPKSFIACTNANDVPR-----FLNNG 288
QY 203 NGDVEIILNADGALDKYLLNNLEPYPTDLMAGOEAHVYKYVARSOLFQCOISITIKEPN 262
DB 289 TWNPKTIYTSINAMD-----ISQPNWRIIEELFYKKMDLKKRGVSVDHTEETL 342
QY 263 SECVRPQCEPQGFQAVKGTGAAPAAQRLILKRSAPENIIDVTD-----INTL 317
DB 343 KELFKL-----GYSEPHAAIAYRLLRDQKLENRGELFLGTAHPAKKNTV 388
QY 318 E-ISDDNALPYDLRHR---ALLOHNGQPV 343
DB 389 EKILNKISLPSELNRIIDPLLSHNTINPV 418

RESULT 11
CGEL_CHICK STANDARD; PRT; 407 AA.
AC P49707; Q91032;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin E1.
GN CCNE1 OR CCNE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC Li H., Lahti J.M., Valentine M., Houston J., Kidd V.J.;
RT "The cyclin E gene is apparently essential in avian B cells.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH A MEMBER OF THE CDK2/CDK PROTEIN KINASES
CC TO FORM A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN
CC SUBUNIT IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX (BY
```

```

CC -1- SIMILARITY.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- PTM: PHOSPHORYLATION BY CDK2 TRIGGERS ITS RELEASE FROM CDK2 AND
CC DEGRADATION VIA THE UBIQUITIN PROTEASOME PATHWAY (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN E SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28981; AAA74981.1; -
DR EMBL: U28990; AAA81647.1; -
DR EMBL: U28982; AAA81647.1; JOINED.
DR EMBL: U28983; AAA81647.1; JOINED.
DR EMBL: U28984; AAA81647.1; JOINED.
DR EMBL: U28985; AAA81647.1; JOINED.
DR EMBL: U28986; AAA81647.1; JOINED.
DR EMBL: U28987; AAA81647.1; JOINED.
DR EMBL: U28988; AAA81647.1; JOINED.
DR EMBL: U28989; AAA81647.1; JOINED.
DR HSSP: P30274; IVIN.
DR InterPro: IPR004366; Cyclin.
DR InterPro: IPR004367; Cyclin_Cterm.
DR Pfam: PF00134; cyclin_1.
DR Pfam: PF02984; cyclin_C; 1.
DR SMART: SM00385; CYCLIN; 1.
DR PROSITE: PS00292; CYCLINS; 1.
DR KMW: Cell cycle; Cell division; Phosphorylation; Nuclear protein.
FT MOD_RES 392 392 PHOSPHORYLATION (By similarity).
FT CONFLICT 106 106 G -> GD (in Ref. 1; AAA81647).
FT CONFLICT 343 343 L -> S (in Ref. 1; AAA81647).
SQ SEQUENCE 407 AA: 46739 MW: A1032C7CB0BC1D1A CRC64;

Query Match
Best Local Similarity 4.3%; Score 86.5; DB 1; Length 407;
Matches 65; Conservative 48; Mismatches 101; Indels 103; Gaps 16;

OY 79 FDSQNAFRTS-----LNPRGIVTTTVISFHPLE----- 109
DB 89 FHLNVASTRSSPLILGMANRDVWKNMINKETVYRDKLQMRHLLQPKRTILLDW 148
OY 110 ---VTVDRARYVQCCEYMEAD-----KTVSAQIEVSEITTFAPQ-TOIYPMRY 152
DB 149 LMEVCEVYKLYR-ETFFLAODEFDRFMAQONVYKTLDLIGISLDFIAKLEIYPPKX 207
OY 153 CRYE-IIDGSPITGQPV--QFAITGQPVYHKTCDSFVDFPCAIV----- 194
DB 208 HQFAVYTDGACTDELLSMELLIMKAL--NMNINPLITVSWLNTYQVAYINLELYVLP 265
OY 195 ---HSCFVDGNGDVEIILNADGALDKYLLNNLEXPDLMAQEAHVYKYADRSQL--- 248
DB 266 QYFQQLFV-----QIAELL--DLCLVD---IGCLEYTGVLAAASALYHFSSELQKVS 315
OY 249 FYQCCQSIITKEPNSRCVPCQSEPGFCAVKTGGAAAPAAAQRLKKSAPENIT 308
DB 316 YEMCEL-----EECVKMWVPPAM--AIREVGS-----KLKHFRGIAPEDILH 355
OY 309 DVRTDINTLEISPDNOA 325
DB 356 NIOTHTINSIDLDKAQA 372

```

```

DE Hypothetical protein C11D3.14c in chromosome I.
GN SPAC11D3.14c.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; Pubmed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer W., Mueller Auer S.,
RA Gabel C., Fuchs M., Filtz C., Holtz E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinrach E., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motter S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga K.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Patahkin J.,
RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.;
RT *The genome sequence of Schizosaccharomyces pombe.*;
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: BELONGS TO THE OXOPROLINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z68166; CAA92315.1; -
DR InterPro: IPR002821; Hydanolase_A.
DR InterPro: IPR003692; Hydanolase_B.
DR Pfam: PF01968; Hydanolase_A; 1.
DR Pfam: PF02538; Hydanolase_B; 1.
DR KMW: Hypothetical protein.
SQ SEQUENCE 1260 AA: 138784 MW: 44F14F0638224BBB CRC64;

Query Match
Best Local Similarity 4.3%; Score 86.5; DB 1; Length 1260;
Matches 73; Conservative 49; Mismatches 111; Indels 153; Gaps 18;

OY 34 GPTSIITNNTNNAEGHYVYK-----GLYDQGCGRN---DEGGRQVAGISLPDSCN 83
DB 856 GITPGSMPSNSKAITEEGCAITFFVYKAGTDEGLQLFLDFEPAK-----YPPCS 907
OY 84 VARTSLN-----PRGIFVTTTVISFHPLEFYTKYVDRAVYVOCFYWEADKTVSA 132
DB 908 GSRTLNDNISDVKAMLSACHRSRSMVEKLYVEYGLDIYOR--SMYGID-----AAA 956
OY 133 QIEVSEITTAFTQTQVPEVPCRYEILDCG-----PTGQPV-QFAITGQPVYHKW 180
DB 957 EKAVRDVILKAFSVQNSQRPFLKAIIDYMDGTPLQLBVKIDIPETGDAVFEGEGPVEYCNW 1016
OY 181 TCDSTVTFPCAIVH-----SCFVD-----DGNQDT 206

```



Db 1017 --NAPAIITSSVYICRSIIINODIPLNEGKLPKIEIRIPSPCLNPFSETAAYGVGNLT 1074  
QY 207 VE-----ILNADG-CALDKYLLNNLEYPTDLMAGOEAHVYKYADRSOLFQCOISITIKE 260  
Db 1075 SQRITDYLKAFSICASOGCMNNLTFCYDGENEE----- 1110  
QY 261 PNSECVBPQSEPGFAGVKT--GGAAPK-----AAAAQLRLKRSAPENIIDVRTDI 314  
Db 1111 -----GFAMYEITAGAGAGPTWNGTSGVHTMTWTRITDPE--VYERRA-- 1153  
QY 315 NTEISDNOALPYDLKRLALQHNG 340  
Db 1154 -----PVILRRPFLRENSG 1167

RESULT 13  
ISCS\_METTE STANDARD; PRT; 404 AA.  
AC P57795;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable cysteine desulfurase (EC 4.4.1.-) (Nifs protein homolog).  
GN ISCS OR NIFS.  
OS Methanosarcina thermophila.  
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20389605; PubMed=10930739;  
RA Borupa B., Ferryb J.G.;  
RT "Cysteine biosynthesis in the archaea: Methanosarcina thermophila  
utilizes O-acetylserine sulfhydrylase.";  
RT FEMS Microbiol. Lett. 189:205-210(2000).  
CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR FROM CYSTEINE  
TO PRODUCE ALANINE (BY SIMILARITY).  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.  
CC -1- CAUTION: The conserved pyridoxal-binding site Lys at position 216  
is replaced by a Glu.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF276772; AGO1802.1; -  
DR InterPro: IPR000192; AminoTransfV.  
DR Pfam: PF00266; aminoTran\_5; 1.  
DR PROSITE: PS00595; AA\_TRANSFER\_CLASS\_5; FALSE\_NEG.  
KW Lyase; Pyridoxal phosphate. BY SIMILARITY.  
FT ACT\_SITE 338  
FT SEQUENCE 404 AA; 44335 MW; 87B5BA25F87E2A25 CRC64;  
SO

Query Match 4.2%; Score 86; DB 1; Length 404;  
Best Local Similarity 21.1%; Pred. No. 4.5;  
Matches 71; Conservative 37; Mismatches 109; Indels 120; Gaps 16;

QY 123 YMEADKTVSAQIEYSETTAFQGIYVMPACRYEI----- 157  
Db 19 YMNSATTPVKEVYEMLPYMTENFGNPSIYEIGRTSKHAINARKKADALGAENE 78  
QY 158 ---LDGSPGTG---QPVQFA-----IIGQPYVHKWTCDSFTVDFECAYVHSCFVDDG 202  
Db 79 IYTFSGTESDNNAIKIAIARNRKGHIIITSSIEHH-----AVLHTCAWLEEG 126  
QY 203 NGDTVEILAND--GCALDKYLLNNLEYPTDLMA-----GOEAHVY 240

Db 127 QGFVEVTLPVDRKGVNSPDELRNAIRDDTLLISIMFANNEIGTIOPIKEIGELAKENOIY 186  
QY 241 KYADRSOLFVYCOQISITIKEPNSRECVQSCSE-----POEFGA-----YKT-----GGAA 286  
Db 187 FHTDAVQAI--GHVIDKIKINIDLLSLGHEFEGPKCGALYRKGVKIDPLHGAQE 244  
QY 287 KPAANAQLRLKRSAPENI-----IDVRDINTLESDDNOALPYDLRRRL----- 335  
Db 245 -----RKRAGTEVPGIYGLKATLAEIETSENNRTL-LKLRRLIEGLIKI 292

QY 336 --LOHNGQVYLAAYONGICSPFGSKMGLSTALI 370  
Db 293 PXTHLNGHPT--QRLANNVNT---FXYIEGESILL 324

RESULT 14  
DSG2\_HUMAN STANDARD; PRT; 1117 AA.  
AC Q14126;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Desmoglein 2 precursor (HDGC).  
GN DSG2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Colon carcinoma;  
RC MEDLINE=94192736; PubMed=8143788;  
RA Schaefer S., Koch P.J., Franke W.W.;  
RT "Identification of the ubiquitous human desmoglein, Dsg2, and the  
RT expression catalogue of the desmoglein subfamily of desmosomal  
RT cadherins.";  
RT Exp. Cell Res. 211:391-399(1994).  
RN [2]  
RP SEQUENCE OF 777-1117 FROM N.A.  
RX MEDLINE=92037656; PubMed=1935985;  
RA Koch P.J., Goldschmidt M.D., Walsh M.J., Zimbelmann R., Franke W.W.;  
RT "Complete amino acid sequence of the epidermal desmoglein precursor  
RT polypeptide and identification of a second type of desmoglein gene.";  
RT Eur. J. Cell Biol. 55:200-208(1991).  
CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOSOME JUNCTIONS.  
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE  
CC FILAMENTS MEDIATING CELL-CELL ADHESION.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -1- TISSUE SPECIFICITY: ALL OF THE TISSUES TESTED AND CARCINOMAS.  
CC -1- DOMAIN: CALCINUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS  
(POTENTIAL).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z26317; CAA81226.1; -  
DR HSPG; P1S116; INCT.  
DR Genew; HGNC:3049; DSG2.  
DR KIM; 125671; -  
DR InterPro: IPR002126; Cadherin.  
DR Pfam: PF00028; cadherin; 4.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA\_4.  
DR PROSITE: PS00232; CADHERIN\_1; 3.  
DR PROSITE: PS50268; CADHERIN\_2; 4.  
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;  
KW Cytoskeleton; Calcium-binding.



```
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 48 POTENTIAL.
FT CHAIN 49 1117 DESMOGLEIN 2.
FT DOMAIN 49 608 EXTRACELLULAR (POTENTIAL).
FT TRASMEM 609 633 POTENTIAL.
FT DOMAIN 634 1117 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 49 159 CADHERIN 1.
FT DOMAIN 160 272 CADHERIN 2.
FT DOMAIN 273 387 CADHERIN 3.
FT DOMAIN 388 502 CADHERIN 4.
FT REPEAT 880 911 DESMOGLEIN REPEAT 1.
FT REPEAT 912 941 DESMOGLEIN REPEAT 2.
FT REPEAT 942 967 DESMOGLEIN REPEAT 3.
FT REPEAT 968 991 DESMOGLEIN REPEAT 4.
FT REPEAT 992 1020 DESMOGLEIN REPEAT 5.
FT REPEAT 1021 1050 DESMOGLEIN REPEAT 6.
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1117 AA; 122385 MW; 2238897PBD70B289 CRC64;
```

Query Match 4.2%; Score 85.5; DB 1; Length 1117;  
Best Local Similarity 25.6%; Pred. No. 20;  
Matches 52; Conservative 27; Mismatches 81; Indels 43; Gaps 12;

```
OY 15 LSYSPVDNGVEGEPEI--ECGPTSTINFTNRNAFEG-----HYVKGLYDQEGOR 64
DB 347 LDFSVYVANKAKAFHKSIRKRYPTPIPIVKKVKNKEGJHFHSSVSTIVSESMD---R 402
OY 65 NDEGROVAGISLPDSCN---VAFTRSLNPRGIF---VTTVVISFHLFVTVK-- 113
DB 403 SSXG--QIIGNQAEFDEDTGLPAHARYVLEDRDNMISVSVTSEIKLKPDESRYVQ 460
OY 114 DDAYRYOCYMEAD---KTVSAQ--IEVSEITTAFOQVPM-PCRC-----YEIID 159
DB 461 NQTYVTKYVAISEDYPRKTIYTGTVLVLENDINDNCPTLIEPVOTICHAERYVNTAEDLD 520
OY 160 GPGTGPVQFAIIGQP--YVHKW 180
DB 521 GHPNSGPFSEFVIDKPPGMAEKW 543
```

```
RESULT 15
HIRA_FUGRU
ID HIRA_FUGRU STANDARD; PRT; 1025 AA.
AC 042611;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE HIRA protein (Tup1 like enhancer of split protein 1).
GN HIRA OR TUPLE1.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
ON NCB1_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.,
RX MEDLINE=98201624; PubMed=9524281;
RA Llavador R., Estivill X., Scambler P., Pritchard M.;
RT "Isolation and genomic characterization of the TUPLE1/HIRA gene of
the pufferfish Fugu rubripes.";
RL Gene 208;779-283(1998).
CC -1- FUNCTION: Could have a part in mechanisms of transcriptional
regulation similar to that played by yeast HIR1 and HIR2 together.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE HIR1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; U94325; AAC60370.1; -.
CC EMBL; U94324; AAC60369.1; -.
CC InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 8.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PRODOM; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 3.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repeat; WD repeat; Nuclear protein.
FT REPEAT 68 98 WD 1.
FT REPEAT 129 159 WD 2.
FT REPEAT 172 202 WD 3.
FT REPEAT 266 313 WD 4.
FT DOMAIN 673 682 POLY-ALA.
FT DOMAIN 685 688 POLY-ALA.
SQ SEQUENCE 1025 AA; 111856 MW; A4212152D75B6A37 CRC64;
```

Query Match 4.2%; Score 85; DB 1; Length 1025;  
Best Local Similarity 22.1%; Pred. No. 19;  
Matches 44; Conservative 37; Mismatches 80; Indels 38; Gaps 8;

```
OY 200 DDGNDGYEILNADGALDKYLLNLEYPPTDLMAGQEAHVYVYADRSQLFYQCQISTITK 259
DB 604 EDTSSDSEDKM---ATINKNLAFNKKRPELLMDGAEVYERKKKRRP---KDMAASIA 656
OY 260 EPNSECVRCQCEPQGFQAVKGTGAAPAAQRL---LKKR-----SAPENIIDV 310
DB 657 QPLTYTTSFAEREPS--RAAAGAGAAAPTAALKLPTPSIKKFTLYOVSDPSVLEV 714
OY 311 PFDINTLEIS-----DDNQALPYDLRHRALLQHNQOPVITLAVNGICMSPF- 357
DB 715 ENEVSVAAGSRSLQRCRSDGRDQNTLLP-----SSVLTAAGSSDVAVASDRRLSVFS 769
OY 358 --GFSEMGSLSTALIAAVI 374
DB 770 SCGRRLPALQLATPASPAL 788
```

```
RESULT 16
DYHG_CHLRE
ID DYHG_CHLRE STANDARD; PRT; 4485 AA.
AC 039575;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dynein gamma chain, flagellar outer arm.
GN ODA2 OR ODA-2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
ON NCB1_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=1132D;
RA MEDLINE=94274766; PubMed=7516341;
RA Wilkerson C.G., King S.W., Wiltman G.B.;
RT "Molecular analysis of the gamma heavy chain of Chlamydomonas
flagellar outer-arm dynein.";
RL J. Cell Sci. 107:497-506(1994).
CC -1- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CLIA AND
FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.
CC DYNEIN HAS ATPASE ACTIVITY.
CC -1- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND
GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.
CC -----
```

CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U15303; AAA50455.1; -.  
 CC Interpro: IPR004273; Dynein\_heavy.  
 CC Interpro: IPR001230; Ptenyl\_site.  
 CC Pfam: PF03028; Dynein\_heavy; 1.  
 CC Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 CC Coiled coil.  
 KW Coiled coil.  
 FT DOMAIN 449 469 COILED COIL (POTENTIAL).  
 FT DOMAIN 804 838 COILED COIL (POTENTIAL).  
 FT DOMAIN 1093 1114 COILED COIL (POTENTIAL).  
 FT DOMAIN 1275 1297 COILED COIL (POTENTIAL).  
 FT DOMAIN 1699 1727 COILED COIL (POTENTIAL).  
 FT DOMAIN 1917 1945 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 3077 3099 COILED COIL (POTENTIAL).  
 FT DOMAIN 3196 3227 COILED COIL (POTENTIAL).  
 FT DOMAIN 3265 3343 COILED COIL (POTENTIAL).  
 FT DOMAIN 3569 3663 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1819 1826 ATP (POTENTIAL).  
 FT NP\_BIND 2099 2106 ATP (POTENTIAL).  
 FT NP\_BIND 2432 2432 ATP (POTENTIAL).  
 FT NP\_BIND 2802 2809 ATP (POTENTIAL).  
 SO SEQUENCE 4485 AA; 512836 MW; 974B79328B403677 CRC64;  
 Query Match 4.2%; Score 85; DB 1; Length 4485;  
 Best Local Similarity 20.8%; Pred. No. 1,4e+02;  
 Matches 84; Conservative 51; Mismatches 131; Indels 138; Gaps 24;  
 QY 1 MMIRLAF-CCT- LIAL-SXIPVNGVEGEPELE-CGPT-SITINEN 43  
 Db 2771 MKLELVFTDCVTHMTRILAMPGLGLVGVGSGKSLRSATVAGTFYITTKYN 2830  
 QY 44 TRNAFEGHYVYKGLYDEGCR-NDEG-GROVAGISLPEF 80  
 Db 2831 VSNLFE-HIKGLYKTAGFQGVYFLTDAEVKDEGFLEYINQIMTGEVAGILTRKD 2887  
 QY 81 SCVAFRRSLNP-RGIFTVTVYISHPLFVTKV-DRAVVOCFYEAADKYVSAQ 133  
 Db 2888 QDMT-VNDRPVNKHQAPGLDLYDNLNPF-FLNKRVDNLHVLCF- 2931  
 QY 134 IEVSEITTAFOCTIVPMPVCRXYELDGPFGQPVQFALIGQPVYHKMTCDSEVDFCAV 193  
 Db 2932 -SPVAKAFARRAQFP-GING-C-TIDWFCP- 2959  
 QY 194 VHSCEVDGNGDVEIINADGALDKYL-NNLEYPTDLMAQEAHYVYKADRS 246  
 Db 2960 -GPKRLTSSVSGKFDKFTMACPKKVKNOLE-LIMG-HAHVFTAAK 3004  
 QY 247 OLF-YOCOSITTKENSEGVRCSEPOSGFVKTGAAGAARAAQAQELKKBSAEP 304  
 Db 3005 EYFEKIRYVYVYPKS-YLSFLQYKELVYAKKVSFTKELVLEVACOKMEP 3056  
 QY 305 ENIDVPTDINTL-EISDNOALPVDLRHRLALOHNGOPVILA 346  
 Db 3057 -KADVNMKAELAVNKGATVSAKKEAALLKQISESTAI 3094  
 RESULT 17  
 ID CAOC CANFA STANDARD; PRT; 410 AA.  
 AC P12637;  
 DF 01-OCT-1989 (Rel. 12, Created)  
 DF 01-OCT-1989 (Rel. 12, Last sequence update)  
 DF 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Calsequesterin, cardiac muscle isoform precursor.

GN CASQ2.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=8243763; PubMed=3379055;  
 RA Scott B.T., Sliemersman H.K.B., Collins J.H., Nadal-Glunard B.,  
 RA Jones L.R.;  
 RT "Complete amino acid sequence of canine cardiac calsequestrin deduced  
 RT by cDNA cloning.";  
 RL J. Biol. Chem. 263:8958-8964(1988).  
 RN [2]  
 RP PARTIAL SEQUENCE OF 20-71.  
 RC TISSUE=Heart;  
 RX MEDLINE=88107564; PubMed=3427023;  
 RA Slupsky J.R., Ohnishi M., Carpenter M.R., Reithmeier R.A.F.;  
 RT "Characterization of cardiac calsequestrin.";  
 RL Biochemistry 26:6539-6544(1987).  
 RN [3]  
 RP PHOSPHORYLATION BY CK2.  
 RX MEDLINE=91093153; PubMed=1985907;  
 RA Cala S.E., Jones L.R.;  
 RT "Phosphorylation of cardiac and skeletal muscle calsequestrin  
 RT isoforms by casein kinase II. Demonstration of a cluster of unique  
 RT rapidly phosphorylated sites in cardiac calsequestrin.";  
 RL J. Biol. Chem. 266:391-398(1991).  
 CC -1- FUNCTION: CALSEQUESTERIN IS A HIGH-CAPACITY, MODERATE AFFINITY,  
 CC CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE  
 CC IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTERIN THROUGH  
 CC A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40  
 CC TO 50 MOLES OF CALCIUM.  
 CC -1- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTERIN OCCURS IN THE  
 CC SARCOPLASMIC RETICULUM/S TERMINAL CISTERNAE LUMINAL SPACES OF  
 CC CARDIAC AND SLOW SKELETAL MUSCLE CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE CALSEQUESTERIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J03766; AAA30833.1; -.  
 CC PIR: A28071; A28071.  
 CC PIR: A39040; A39040.  
 CC PIR: C27499; C27499.  
 CC HSP: P07221; IABY.  
 CC Interpro: IPR001393; Calsequestrin.  
 CC Pfam: PF01216; Calsequestrin; 1.  
 CC PRINTS: PS00863; CALSEQUESTERIN.  
 CC PROSITE: PS00864; CALSEQUESTERIN\_2; 1.  
 CC PROSITE: PS00864; CALSEQUESTERIN\_2; 1.  
 CC Muscle; Glycoprotein; Calcium-binding; Signal; Phosphorylation.  
 KW SIGNAL 19  
 FT CHAIN 20 410 CALSEQUESTERIN, CARDIAC MUSCLE ISOFORM.  
 FT DOMAIN 219 254 PRO-RICH.  
 FT DOMAIN 221 242 CALCIUM REGULATED HYDROPHOBIC SITE.  
 FT DOMAIN 372 410 ASP/GLU-RICH (ACIDIC).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MOD\_RES 395 395 PHOSPHORYLATION (BY CK2).  
 FT MOD\_RES 397 397 PHOSPHORYLATION (BY CK2).  
 FT MOD\_RES 401 401 PHOSPHORYLATION (BY CK2).  
 FT MOD\_RES 405 405 PHOSPHORYLATION (BY CK2).  
 FT CONFLICT 71 71 Q -> E (IN REF. 2).  
 SO SEQUENCE 410 AA; 47416 MW; FCA99AB07E7AB8B2 CRC64;  
 Query Match 4.2%; Score 84.5; DB 1; Length 410;  
 Best Local Similarity 20.3%; Pred. No. 6.2;



Db 960 AAGFSEP-GDGGITLRCHEGVIGLVTMGG 987

## RESULT 19

MOCS\_BOVIN STANDARD; PRT; 563 AA.

AC P98091;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Submaxillary mucin-like protein.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 ON NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Submaxillary gland;  
 RX MEDLINE=90370871; PubMed=2204065;  
 RA Bhargava A.K., Moltach J.T., Davidson E.A., Bhavanandan V.P.;  
 RT "Cloning and cDNA sequence of a bovine submaxillary gland mucin-like  
 protein containing two distinct domains."  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6798-6802(1990).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: SUBMAXILLARY MUCOSAE.  
 CC -1- SIMILARITY: TO PORCINE APOMUCIN.  
 CC -1- SIMILARITY: CONTAINS 1 WFC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M36192; AAA30657.1; -  
 DR InterPro: IPR000359; Cys\_knot.  
 DR InterPro: IPR001007; WFC\_C.  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR SMART: SM00041; CT; 1.  
 DR SMART: SM00214; WFC; 1.  
 DR PROSITE: PS01208; WFC; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 KW Repeat; Glycoprotein.  
 FT DOMAIN 61 158 3 X 11 AA REPEATS.  
 FT REPEAT 61 71 1.  
 FT REPEAT 112 122 2.  
 FT REPEAT 148 158 3.  
 FT DOMAIN 338 404 WFC.  
 FT DOMAIN 471 555 CTCK.  
 FT DOMAIN 1 282 TO PORCINE APOMUCIN.  
 FT DISULFID 471 518 BY SIMILARITY.  
 FT DISULFID 485 532 BY SIMILARITY.  
 FT DISULFID 494 548 BY SIMILARITY.  
 FT DISULFID 498 550 BY SIMILARITY.  
 FT DISULFID ? 554 BY SIMILARITY.  
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 563 AA; 58913 MW; AB326CD78E5FFCF CRC64;

Query Match 4.1%; Score 83.5; DB 1; Length 563;  
 Best Local Similarity 20.3%; Pred. No. 12;

Matches 61; Conservative 33; Mismatches 115; Indels 91; Gaps 15;

QY 21 VDNQGEPELEICGPTISITINENR-NAFEGHYVYKGLYDQEGGRNDGGRVAGISLFF 79  
 Db 192 IGTGITGTGS---GTSSPGFNAEATTFKEHVHTETTRILSGTTRGRSGTTF---IP- 243

QY 80 DSCNVARBSLNPRIQFVTTTVISFHPLEVT----- 111

Db 244 ESSNTGTSTGVGR-----TSTAIVSGRYGVSESSSPGTSKEASETTGTGISTGTSK 299

QY 112 --KVDRAYRVOCFMEADKTVSAQIEVSEITTAFOIIVPWCYREIILDGPGPOVOF 169

Db 300 SNRTTSSRLP-YPETTVATGEOE-TEKRGCTTSLPPPPACY-----GPIGEKKSP 350

QY 170 AIIQOPYHKWTC-DSETVD---TFCAVHSCFVDDGNGDTVEILNADGALDKY----- 220

Db 351 GDITFANCHKCTCYDAETVCKLEKCPSPCKPDEE--RLVKKKMDTCCEIAYCEPRT 407

QY 221 -LNLNLEPDTDLMAQGEAHYKVKADRSQIFVQGISITIKPENECVRCQSEPGRGAV 279

Db 408 CLENNNDY-----EVGASFADPKNPKCISYSCHN-TGEVAV 441

## RESULT 20

TKT\_BACHD STANDARD; PRT; 666 AA.

AC Q9KAD7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transketolase (EC 2.2.1.1) (TK).  
 GN TKT OR BH2352.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 ON NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde  
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AP001515; BAB06071.1; -  
 DR HSSP: P23254; TTK.  
 DR InterPro: IPR000360; Transketolase.  
 DR Pfam: PF00456; transketolase; 1.  
 DR Pfam: PF02779; transket-1; 1.  
 DR Pfam: PF02780; transketolase\_C; 1.  
 DR TIGRFAMs: TIGR00232; tktase\_bact; 1.  
 DR PROSITE: PS00801; TRANSKETOLASE\_1; 1.  
 DR PROSITE: PS00802; TRANSKETOLASE\_2; 1.  
 KW Transferrase; Thiamine pyrophosphate; Complete proteome.  
 SQ SEQUENCE 666 AA; 72112 MW; EF6179FAF3B84D06 CRC64;

Query Match 4.1%; Score 83; DB 1; Length 666;  
 Best Local Similarity 23.5%; Pred. No. 16;

Matches 39; Conservative 29; Mismatches 66; Indels 32; Gaps 7;

QY 161 GPTGQPVQALITIGQPVYHKWTCDSFTVDFCAVHSCFVDDGNGDTVEIILNADGALDKY 220  
 Db 116 GILGQGVAMA-VGMAMERHILATYNRDGYNIVDHYTYTICGDGDLMEGVASABAASLAGH 174

```

OY 221 L-----LNNLEYPTDL-----MAGQEAHYKYKADRSQFYQCOISITIK 255
DB 175 LKGRMLLYDSNDISIDGDHHSFSESVEDREKAYGMHWVRKEDNNL---DEIANAIE 231
OY 260 EPNSECVPOCSEPO---GFGAVTGGGAAK---PAAAOULRLKK 299
DB 232 EAKAD-ERPSELEKTTIGFGSPRKGGKSVSHGAPLGADEVKLKE 276

RESULT 21
ID PD21_ARATH STANDARD: PRT: 820 AA.
AC P58766:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Phospholipase D zeta (EC 3.1.4.4) (AtPLDzeta) (PLD zeta).
GN PLDZETA OR AT5G5370 OR F18618.110
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneo T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Nanno K., Okumura S., Shinpo S., Takeuchi C., Mada T.,
RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston J.,
RA Haberman K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Storeking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haekensen B., Lamar E., Latelle P.,
RA Leonard S., Meyer R., Milvaney E., Ozeresky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeikum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Maria M.,
RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entlan K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,
RA Ranspenger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirke W., Moolman P., Klein Lankhorst R.,
RA Weitzenegger T., Bothe G., Rose M., Hauf J., Bernerstein S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielens J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.,
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
thaliana."
RL Nature 408:823-826(2000).
CC -1- FUNCTION: Hydrolyzes glycerol-phospholipids at the terminal
phosphodiesteric bond.
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O -> choline + a
phosphatidate.
CC -1- COFACTOR: Calcium (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
similarity).
CC -1- DOMAIN: C2 domain is a calcium-binding fold, and the binding
promotes the protein association with membranes. A lower affinity
toward calcium can be anticipated for PLD alpha due to the absence
of two potential calcium ligands.
CC -1- SIMILARITY: Belongs to the phospholipase D family. C2-PLD
subfamily.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PLD PHOSPHODIESTERASE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
CC -----
DR EMBL: AC006258; NOT_ANNOTATED_CDS.
DR PROSITE: PS50004; C2_DOMAIN_2; FALSE_NEG.
DR PROSITE: PS50035; PLD; 2.
KW Hydrolyase; Lipid degradation; C2 DOMAIN.
FT DOMAIN 1 117
FT DOMAIN 334 371
FT DOMAIN 662 689
FT ACT_SITE 339 359
FT ACT_SITE 341 341
FT ACT_SITE 346 346
FT ACT_SITE 667 667
FT ACT_SITE 669 669
FT ACT_SITE 674 674
FT ACT_SITE 674 674
SQ SEQUENCE 820 AA; 93362 MW; FFB9247276CBB7B CRC64;

Query Match 4.1%; Score 83; DB 1; Length 820;
Best local Similarity 22.6%; Pred. No. 22;
Matches 72; Conservative 48; Mismatches 102; Indels 96; Gaps 22;

OY 54 VKGLYDDEGRNDEGGROYAGISLPPDSCNVART-RSLNPRGIFVTTVVISFHPLEVTK 112
DB 46 IKRLTD--SCTSLDFGHLVATIDL--DRSRVARTMRMRPKML-----QSFHYTANS 94
OY 113 VDRAYVQCFYMAADTVGAQ-----IVSETTAFQOIQYPMPCRT-ELIDGPTQ 165
DB 95 ISKT---IFTYKEDPVSASLIGRAYLVTEYITG-----QIDRWLDLD--ENRR 141
OY 166 PVOFAIIGQPVYH---KWTGSETVDTFCVAVHSCFVDGNDPTVEILN-ADGAL----- 217
DB 142 PIQ-----GSKLHVRKFKFHYQDVVMKNGIILPSF---NGVPNAYFNQREGKVTYXQ 193
OY 218 DKYLLNLEYPTDLMAQEAHYKYTA-----DRSQFYQCOISIT-----IKEP 261
DB 194 DAHVIL--EYPPVTLTGGOV-IYKHRCWEIEITDAIWEAKHLIYIAGWSVNTDVTLYRDP 250
OY 262 NSRCVAPQCSSEPOGFAVVTGGAAKPAQAOLRLKKRSAPENIDVPTINTLEISD 321
DB 251 KR--TRPG-----GDLKLG-----ELKKRAENVTVL-----MLVWD 281
OY 322 DNQALPVDLRHRLALQHN 339
DB 282 DRTSHVEFKRDMGMTHD 299

RESULT 22
PAPA_HUMAN STANDARD: PRT: 1627 AA.
ID PAPA_HUMAN 013319; 008371; 09UDK7;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pregnancy-associated plasma protein-A precursor (EC 3.4.24.-) (PAP-A)
DE (Insulin-like growth factor-dependent IGF binding protein-4 protease)
DE (IGF-dependent IGFBP-4 protease) (IGFBP-4ase).
GN PAPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RC TISSUE=Placenta;
RX MEDLINE=96203921; PubMed=8620868;
RA Haaning J., Oxvig C., Overgaard M.T., Ebbesen P., Kristensen T.,
RA Sottrup-Jensen L.;
RT "Complete cDNA sequence of the preproform of human pregnancy-
associated plasma protein-A. Evidence for expression in the brain and
induction by CAMP."

```

RL Eur. J. Biochem. 237:159-163(1996).  
 RN [12]  
 RP SEQUENCE OF 77-1627 FROM N.A., SEQUENCE OF 81-98; 117-126; 210-224;  
 RP 466-485; 507-519; 576-593; 609-621; 718-736; 742-754; 1006-1017;  
 RP 1357-1273; 1369-1374; 1389-1398; 1490-1509; 1524-1533 AND 1537-1544,  
 RP VARIANT SER-944, AND TISSUE SPECIFICITY.  
 RC TISSUE-SPECIFICITY, AND SERUM;  
 RX MEDLINE=94146014; PubMed=7508748;  
 RA Kristensen T., Oxvig C., Sand O., Moller N.P.H., Sottrup-Jensen L.;  
 RT "amino acid sequence of human pregnancy-associated plasma protein-A  
 RT derived from cloned cDNA.";  
 RL Biochemistry 33:1592-1598(1994).  
 RN [13]  
 RP SEQUENCE OF 81-89; 117-126; 210-224; 460-485; 507-519; 576-593;  
 RP 718-736; 742-754; 1259-1273; 1369-1374; 1490-1509; 1524-1533 AND  
 RP 1537-1544, SUBUNITS, AND INTERCHAIN DISULFIDE BOND.  
 RC TISSUE-SERUM;  
 RX MEDLINE=93286045; PubMed=7685339;  
 RA Oxvig C., Sand O., Kristensen T., Gleich G.J., Sottrup-Jensen L.;  
 RT "Circulating human pregnancy-associated plasma protein-A is disulfide-  
 RT bridged to the proform of eosinophil major basic protein.";  
 RL J. Biol. Chem. 268:12243-12246(1993).  
 RN [14]  
 RP IDENTIFICATION, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE-Fibroblast;  
 RX MEDLINE=91179030; PubMed=10077652;  
 RA Lawrence J.B., Oxvig C., Overgaard M.T., Sottrup-Jensen L.,  
 RA Gleich G.J., Hays L.G., Yates J.R. III, Conover C.A.;  
 RT "The insulin-like growth factor (IGF)-III-dependent IGF binding protein-4  
 RT protease secreted by human fibroblasts is pregnancy-associated plasma  
 RT protein-A.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:3149-3153(1999).  
 RN [15]  
 RP FUNCTION, SUBUNITS, AND ENZYME REGULATION.  
 RX MEDLINE=20469470; PubMed=10913121;  
 RA Overgaard M.T., Haaning J., Boldt H.B., Olsen I.M., Laursen L.S.,  
 RA Christiansen M., Gleich G.J., Sottrup-Jensen L., Conover C.A.,  
 RA Oxvig C.;  
 RT "Expression of recombinant human pregnancy-associated plasma protein-A  
 RT and identification of the proform of eosinophil major basic protein  
 RT as its physiological inhibitor.";  
 RL J. Biol. Chem. 275:31128-31133(2000).  
 RN [16]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=95057018; PubMed=7526035;  
 RA Bono M., Oxvig C., Kephart G.M., Wagner J.M., Kristensen T.,  
 RA Sottrup-Jensen L., Gleich G.J.;  
 RT "Localization of pregnancy-associated plasma protein-A and  
 RT colocalization of pregnancy-associated plasma protein-A messenger  
 RT ribonucleic acid and eosinophil granule major basic protein messenger  
 RT ribonucleic acid in placenta.";  
 RL Lab. Invest. 71:560-566(1994).  
 RN [17]  
 RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
 RX MEDLINE=99423540; PubMed=10491647;  
 RA Overgaard M.T., Oxvig C., Christiansen M., Lawrence J.B.,  
 RA Conover C.A., Gleich G.J., Sottrup-Jensen L., Haaning J.,  
 RA "Messenger ribonucleic acid levels of pregnancy-associated plasma  
 RT protein-A and the proform of eosinophil major basic protein;  
 RT expression in human reproductive and nonreproductive tissues.";  
 RL Biol. Reprod. 61:1083-1089(1999).  
 RN [8]  
 RP DEVELOPMENTAL STAGE.  
 RX MEDLINE=95293954; PubMed=7539791;  
 RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,  
 RA Sigbrand T., Gleich G.J., Sottrup-Jensen L.;  
 RT "Identification of angiotensinogen and complement C3d as novel  
 RT proteins binding the proform of eosinophil major basic protein in  
 RT human pregnancy serum and plasma";  
 RL J. Biol. Chem. 270:13645-13651(1995).  
 CC -1- FUNCTION: Metalloproteinase which specifically cleaves IGFBP-4 in  
 CC the presence of IGF, resulting in release of bound IGF.

CC -1- ENZYME REGULATION: Inhibited by complexation with the proform  
 CC of PRG2.  
 CC -1- SUBUNIT: Homodimer; disulfide-linked. In pregnancy serum,  
 CC predominantly found as a disulfide-linked 2:2 heterotetramer with  
 CC the proform of PRG.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: High levels in placenta and pregnancy serum.  
 CC In placenta, expressed in x cells in septa and anchoring villi,  
 CC and in syncytiotrophoblasts in the chorionic villi. Lower levels  
 CC are found in a variety of other tissues including kidney,  
 CC myometrium, endometrium, ovaries, breast, prostate, bone marrow,  
 CC colon, fibroblasts and osteoblasts.  
 CC -1- DEVELOPMENTAL STAGE: Present in serum and placenta during  
 CC pregnancy; levels increase throughout pregnancy.  
 CC -1- INDUCTION: By 8-bromodeoxyne-3',5'-phosphate.  
 CC -1- PTM: There appear to be no free cysteinyl groups.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M46.  
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U28727; AAC50543.1; -;  
 DR EMBL: X68280; CAA48341.1; -;  
 DR MEROPS: M46.001; -;  
 DR Genew: HGNC:8602; PAPA.  
 DR MIM: 176385; -;  
 DR Interpro: IPR000800; Notch.  
 DR Interpro: IPR000436; Sushi\_SCR\_CCP.  
 DR Interpro: IPR000130; Zn\_MTpepcase.  
 DR Pfam: PF00084; Sush1; 4.  
 DR SMART: SM00032; CCP; 4.  
 DR SMART: SM00004; NL; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolyase; Metalloproteinase; Metal-binding; Zinc; Signal; Glycoprotein;  
 KW Zymogen; Repeat; Sushi.  
 FT SIGNAL 1 22  
 FT PROPEP 23 80  
 FT CHAIN 81 1627  
 FT DOMAIN 24 83  
 FT DOMAIN 1215 1280  
 FT DOMAIN 1285 1342  
 FT DOMAIN 1346 1410  
 FT DOMAIN 1415 1471  
 FT METAL 562 562  
 FT ACT\_SITE 563 563  
 FT METAL 566 566  
 FT METAL 566 566  
 FT DISULFID 461 461  
 FT CARBOHYD 390 390  
 FT CARBOHYD 402 402  
 FT CARBOHYD 429 429  
 FT CARBOHYD 480 480  
 FT CARBOHYD 601 601  
 FT CARBOHYD 619 619  
 FT CARBOHYD 725 725  
 FT CARBOHYD 825 825  
 FT CARBOHYD 1026 1026  
 FT CARBOHYD 1222 1222  
 FT CARBOHYD 1226 1226  
 FT CARBOHYD 1323 1323  
 FT CARBOHYD 1465 1465  
 FT CARBOHYD 1519 1519  
 FT VARIANT 944 944  
 FT CONFLICT 511 512  
 FT SEQUENCE 1627 AA; 181005 MW; 34164AA3BE56C768 CRC64;  
 SO  
 R -> S.  
 R/FTId=VAR\_011419.  
 TH -> RD (IN REF. 2; AA SEQUENCE).  
 CC -1- FUNCTION: Metalloproteinase which specifically cleaves IGFBP-4 in  
 CC the presence of IGF, resulting in release of bound IGF.







CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE  
 CC N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY  
 CC FOR COMPLEXING SUBUNITS B AND C. THE C-TERMINUS MAY ALSO REGULATE  
 CC THE CATALYTIC ACTIVITIES OF THE ENZYME.  
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. HIGH  
 CC SIMILARITY WITH FUNGAL DNA POLYMERASE II.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: S60080; AAA15448.1; ALT\_SEQ.  
 DR EMBL: L09561; AAC19148.1; -.  
 DR EMBL: U49356; AAA90924.1; -.  
 DR PIR: A46692; A46692.  
 DR Genew: HGNC:9177; POLE.  
 DR MIT: 174762; -; POLE.  
 DR Interpro: IPR002064; DNA\_POL\_B.  
 DR Pfam: PF00136; DNA\_POL\_B.1.  
 DR SMART: SM00486; POLBc.1.  
 KW Transferrase: DNA-directed DNA polymerase; DNA replication; DNA repair;  
 KM DNA-binding; Zinc-finger; Nuclear protein.  
 FT ZN\_FING 2158 2190 C4-TYPE (POTENTIAL).  
 FT ZN\_FING 2221 2238 C4-TYPE (POTENTIAL).  
 SQ SSQDNCE 2286 AA: 261530 MW: A216FELEAD137DEC CRC64;  
 Query Match 4.0%; Score 81.5; DB 1; Length 2286;  
 Best Local Similarity 22.5%; Pred. No. 1.2e+02;  
 Matches 42; Conservative 26; Mismatches 70; Indels 49; Gaps 8;  
 -----  
 QY 153 CRVEILDGPTGPVQFAIIGOPY-----HKMTCDSEVTDFCAVHSCFVD 201  
 DB 1076 CRV-IISRKEGSPVTERAIPFAFOAEPTVRKHFLEKRWLSSLODF----- 1122  
 QY 202 GNGDPTVEILNAD-----GCALDKYL-----LNLEYPTDLAMG-----OEAHYK 241  
 DB 1123 ---DIRAILMDVYIERLGSIOKIIITIPALQOVKNPVPVKHPMDLHKLEKNVYK 1179  
 QY 242 YADRSOLF-YOCQISITIKENSCVRCQCEPGFGVKTGGAAPAAAOERLLKKR 300  
 DB 1180 QKTSLEFTEBGRQVYMAESBDSPPSAPDMEDFGLVLPDPA-PVTVKRRRVLMES 1238  
 QY 301 SAEPEPT 307  
 DB 1239 QEESQDL 1245  
 -----  
 RESULT 25  
 TORA\_ECOLI STANDARD; PRT; 848 AA.  
 ID TORA\_ECOLI  
 AC P33225; P78227;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE trimethylamine-N-oxide reductase 1 precursor (EC 1.6.6.9) (TMAO  
 DE reductase 1) (trimethylamine oxidase 1).  
 GN TORA OR B0997.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 40-46.  
 RC STRAIN-K12;  
 RX MEDLINE=94293785; PubMed=8022286;  
 RA Mejean V., Jobbi-Nivol C., Lepelletier M., Giordano G., Chipaux M.,  
 RA Pascal M.-C.;  
 RT "TMAO anaerobic respiration in Escherichia coli: involvement of the

RT for operon.";  
 RL Mol. Microbiol. 11:1169-1179(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=9742617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:123-147(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Aiba H., Bada T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 12.7-28.0 mln region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 RN [4]  
 RP SEQUENCE OF 767-848 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=94134696; PubMed=8302830;  
 RA Deguchi C., Kakeda M., Yamada H., Mizuno T.;  
 RT "An analogue of the DnaJ molecular chaperone in Escherichia coli.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1054-1058(1994).  
 CC -1- FUNCTION: Reduces trimethylamine-N-oxide (TMAO) into  
 CC trimethylamine; an anaerobic reaction coupled to energy-yielding  
 CC reactions.  
 CC -1- CATALYTIC ACTIVITY: NADH + trimethylamine-N-oxide = NAD(+) +  
 CC trimethylamine + H(2)O.  
 CC -1- COFACTOR: Molybdenum (Molybdopterin) (By similarity).  
 CC -1- SUBUNIT: Interacts with the N-terminal domain of tofC.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING  
 CC OXIDOREDUCTASE FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: X73888; CA52095.1; -.  
 DR EMBL: AE000201; AAC74082.1; -.  
 DR EMBL: D90736; BAA36139.1; -.  
 DR EMBL: D90737; BAA35764.1; -.  
 DR EMBL: D16500; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: S34222; S34222.  
 DR HSSP: O87948; ITMO.  
 DR EcGene; E311814; torA.  
 DR Interpro: IPR001467; Prok\_Mboxred.  
 DR Pfam: PF00384; molybdopterin.1.  
 DR Pfam: PF01568; Molybdop-binding.1.  
 DR TIGRFAMs; TIGR00509; bisco\_fam.1.  
 DR PROSITE; PS00551; MOLYBDOPTERIN\_PROK\_1; FALSE\_NEG.  
 DR PROSITE; PS00490; MOLYBDOPTERIN\_PROK\_2; 1.  
 DR PROSITE; PS00932; MOLYBDOPTERIN\_PROK\_3; 1.  
 KW Oxidoreductase; NAD; Molybdenum; Periplasmic; Signal;  
 KW Complete proteome.  
 FT SIGNAL 1 39  
 FT CHAIN 1 848  
 FT CONFLICT 173 173 TRIMETHYLAMINE-N-OXIDE REDUCTASE 1.  
 FT CONFLICT 176 176 A -> R (IN REF. 1).  
 FT CONFLICT 256 256 A -> R (IN REF. 1).

```

FT CONFLICT 258 258 V -> S (IN REF. 1).
FT CONFLICT 281 281 R -> G (IN REF. 1).
FT CONFLICT 325 325 Q -> E (IN REF. 1).
FT CONFLICT 348 348 Q -> S (IN REF. 1).
FT CONFLICT 503 504 KL -> NV (IN REF. 1).
FT CONFLICT 713 714 QQ -> ME (IN REF. 1).
FT CONFLICT 751 751 P -> L (IN REF. 1).
FT CONFLICT 781 781 P -> L (IN REF. 1 AND 3).
SQ SEQUENCE 848 AA; 94456 MW; 59DDACB00B1843E7 CRC64;

Query Match
Best Local Similarity 4.0%; Score 81; DB 1; Length 848;
Matches 56; Conservative 22; Mismatches 72; Indels 78; Gaps 15;

QY 42 FNTNAAE-----GHVYKGLVD-----OEGCRNDEGGQVAGISLP--FDS 81
DB 558 FEARNDEDFIELCRNREBEAFEGIDEMKMRINGEGVQGGK-----RGVHLPADD 613
QY 82 CNVARTSLNPRGIFVTTTVISF-HP-LEVTKYDRAVRVOCFYMAD---KTVSAQIE 135
DB 614 -----FMNNKEVEFDHPOMFV-----RHQAFREDPLELTPSGLIE 652
QY 136 VSEITTAFCQ---QIYPMPCREIILDGPTGQVQFAIIIGQPVY----HKVTCSET 186
DB 653 IYKTIADMYDDCGHPMFPEKERSHGPGSO--KYPHLQSVHPDFRLHSQLCSET 710
QY 187 VDFECAYV--HSCFVD-----GNQTYEILNADGAL-----DKY 220
DB 711 LRQOYVAKREPVFINPDASARGIRNGDVVRFNANGVLAGVNSDXY 760

RESULT 26
CARB_XYLFA STANDARD; PRT; 1080 AA.
ID CARB_XYLFA
AC Q9PECL1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
GN carbamoyl-phosphate synthase ammonia chain).
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala J.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Btlones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa P.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Pacinencini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Kleiger J.C., Kuramae E.E., Laigret F., Lambdas M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vicorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Neto L.E.S.,
RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Pexoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tshako M.H.,

```

```

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -1- COFACTOR: Binds three manganese ions (by similarity).
CC -1- PATHWAY: Arginine biosynthesis.
CC -1- SUBUNIT: Pyrimidine biosynthesis; first step.
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (by
CC similarity). BELONGS TO THE CARB FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage is free and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL; AE003946; AAF83917.1; -.
CC HSSP; P00968; 1A9X.
DR InterPro: IPR005483; CPase_L.
DR InterPro: IPR005479; CPase_L_D2.
DR InterPro: IPR005480; CPase_L_D3.
DR InterPro: IPR005481; CPase_L_N.
DR InterPro: IPR004362; MGS_Like.
DR Pfam: PF00289; CPase_L_chain; 2.
DR Pfam: PF02786; CPase_L_D2; 2.
DR Pfam: PF02787; CPase_L_D3; 1.
DR Pfam: PF02142; MGS; 1.
DR PRINTS; PR00098; CPASE.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 403 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 404 554 OLIGOMERIZATION DOMAIN.
FT DOMAIN 555 942 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 943 1080 ALLOSTERIC DOMAIN.
FT REPEAT 1 554
FT REPEAT 555 1080
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 303 354 ATP (POTENTIAL).
FT METAL 285 285 MANGANESE 1 (BY SIMILARITY).
FT METAL 299 299 MANGANESE 2 (BY SIMILARITY).
FT METAL 301 301 MANGANESE 3 (BY SIMILARITY).
FT METAL 830 830 MANGANESE 3 (BY SIMILARITY).
FT METAL 847 847 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1080 AA; 117383 MW; D8CED71DD42328B7 CRC64;

Query Match
Best Local Similarity 4.0%; Score 81; DB 1; Length 1080;
Matches 80; Conservative 53; Mismatches 143; Indels 112; Gaps 20;

QY 50 GHVYKGLVD--EGCRNDE-----GROVAGISLPDSCNARTSLNPRGIFVTTT 100
DB 540 GTALIXSTYEECEAPSDRRKIMILGGPNRIGOGIEFYCVHAALALREDF---ET 596
QY 101 VVISFHPLEF-TKYDRAVRVOCFYMADKTVSAQIEVSEI-----TT 141
DB 597 IMVKNPETSVDTDTSRL---YFE-PLTLEDVLEIYEVHPRGVLYVGQGPPLAK 652
QY 142 AFQIQIYPMPCREIILDGCTGQVQFAI---IGQPVHKVTCDESETVDT----- 189
DB 653 ALEANGVPVIGTSPESIDLEDRERFQKLVQGLRQRP---NCTARTEALVAREI 708
QY 190 -FCAYVNSCFDDG-----DYEILNADGALDKYLLNNLEYPDTLMA 233
DB 709 GYPLVVRPSYVLGGRAMEIYVGADLARVYRDAAVKSNDSPVLLDRFLDNAVEVDVIA 768

```





CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC euroids I.; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxId=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Birte;  
 RX MEDLINE=90147555; PubMed=2515655;  
 RA Ealing P.M., Casey R.;  
 RT "The cDNA cloning of a pea (*Pisum sativum*) seed lipoxygenase.  
 RT Sequence comparisons of the two major pea seed lipoxygenase  
 RT isoforms.";  
 RL Biochem. J. 264:929-932(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Birte;  
 RA Casey R.;  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -! FUNCTION: PLANT LIPOXYGENASE MAY BE INVOLVED IN A NUMBER OF  
 CC DIVERSE ASPECTS OF PLANT PHYSIOLOGY INCLUDING GROWTH AND  
 CC DEVELOPMENT, PEST RESISTANCE, AND SENESCENCE OR RESPONSES TO  
 CC WOUNDING. IT CATALYZES THE HYDROPEROXIDATION OF LIPIDS, CONTAINING  
 CC A CIS-, CIS-,4-PENTADIENE STRUCTURE.  
 CC -! CATALYTIC ACTIVITY: linoleate + O(2) = (9Z,11E)-(13S)-13-  
 CC hydroperoxyoctadeca-9,11-dienoate.  
 CC -! COFACTOR: IRON.  
 CC -! SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -! SIMILARITY: BELONGS TO THE LIPOXYGENASE FAMILY.  
 CC  
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -  
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announcement/>  
 CC OR SEND AN EMAIL TO [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC -----  
 DR EMBL; X17061; CA934906.1; -.  
 DR EMBL; X78580; CA935318.1; -.  
 DR PIR; S07075; S07075.  
 DR HSSP; P08170; 25BL.  
 DR InterPro; IPRO00907; Lipoxygenase.  
 DR InterPro; IPRO01024; Lipoxygenase\_LH2.  
 DR Pfam; PF00305; Lipoxygenase\_1.  
 DR Pfam; PF01477; PLAT; 1.  
 DR PRINTS; PRO0087; LIPOXYGENASE.  
 DR SMART; SM00308; LH2; 1.  
 DR PROSITE; PS00081; LIPOXYGENASE\_2; 1.  
 DR PROSITE; PS00711; LIPOXYGENASE\_1; 1.  
 KM Oxidoreductase; Dioxygenase; Iron; Multigene family.  
 FT METAL 524 524  
 FT METAL 529 529 IRON (BY SIMILARITY).  
 FT METAL 716 716 IRON (BY SIMILARITY).  
 FT METAL 864 864 IRON (BY SIMILARITY).  
 FT CONFLICT 333 333 L->I (IN REF. 2).  
 FT CONFLICT 562 562 MISSING (IN REF. 2).  
 FT CONFLICT 578 578 L->I (IN REF. 2).  
 SO SEQUENCE 864 AA; 97133 MW; 2919AF5FAF272CDF CRC64;

Query Match 4.0%; Score 80.5; DB 1; Length 864;  
 Best local Similarity 18.3%; Pred. No. 39;  
 Matches 48; Conservative 29; Mismatches 76; Indels 109; Gaps 9;

OY	36	TSTITNFTNAFAEGHYVVKGLYDEGCRNDGGRGQVAGISLPFDSCWAVRTSLNPRGI	95
Db	298	SADFNTVF-TPNMDFDSFOVRNLFE-----GKIKPLD-----	328
OY	96	FVTTVVISHPDLFYTKVDRAVRYOCFYMEADKTVAISOIEVEITTAFOGIVMPVCRY	155
Db	329	-----VISTLSPLPYVK-----EIFRTDGEV-----LKFTPPHYIRV	361
OY	156	EILDGGPGRGQVRQAIIIGQPVYHKHTCSSEVYDTFCAYVHSQV-----	199
Db	362	S-----KSAMMTDEEFAREMLAGVNPDMCTIRGLOEFPKRSLDAE	401

QY 200 --DDGNGDVEILNADGCAIDKLYLNINLEXPID-----IMAGOPAHVYKADRSOLF 249  
 DB 402 YGDHSTKISVVDVNLGDGCTIDALSGRLFIIDYHDTFIPFLRINETSARAVATRTILF 461  
 QY 250 YO-----COISTIKEPENSE 264  
 DB 462 LKENGTLKRVAIETLSLPHPDGD 483

RESULT 31  
 TORA\_ECO57  
 ID TORA\_ECO57 STANDARD; PRT; 848 AA.  
 AC P58360.  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Trimethylamine-N-oxide reductase 1 precursor (EC 1.6.6.9) (TMAO reductase 1) (Trimethylamine oxidase 1).  
 GN TORA OR Z1415 OR ECS1152.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
 OC NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamasis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"  
 RT Nature 409:529-533(2001).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.;  
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12,"  
 RT DNA Res. 8:11-22(2001).  
 RL -1- FUNCTION: Reduces trimethylamine-N-oxide (TMAO) into trimethylamine; an anaerobic reaction coupled to energy-yielding reactions (By similarity).  
 CC -1- CATALYTIC ACTIVITY: NADH + trimethylamine-N-oxide = NAD(+) + trimethylamine + H(2)O.  
 CC -1- COFACTOR: Molybdenum (Molybdopterin) (By similarity).  
 CC -1- SUBUNIT: Interacts with the N-terminal domain of torC (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING OXIDOREDUCTASE FAMILY.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/or-send-an-email-to-licensese@isb-sib.ch>).  
 CC EMBL: AF005293; AAC5554.1; -  
 DR EMBL: AP002554; BAB34575.1; -  
 DR InterPro: IPR001467; Prok\_Mboxred.  
 DR Pfam: PF00384; Molybdopterin; 1.  
 DR Pfam: PF01568; Molybdopterin; 1.  
 DR TIGRfams: TIGR00509; b1sc\_fam; 1.

DR PROSITE: PS00551; MOLYBDOPTERIN\_PROK\_1; FALSE\_NEG.  
 DR PROSITE: PS00490; MOLYBDOPTERIN\_PROK\_2; 1.  
 DR PROSITE: PS00932; MOLYBDOPTERIN\_PROK\_3; 1.  
 KW Oxidoreductase; NAD; Molybdenum; Periplasmic; signal;  
 KW Complete proteome.  
 FT SIGNAL 1 39 BY SIMILARITY.  
 FT CHAIN 40 848 TRIMETHYLAMINE-N-OXIDE REDUCTASE 1.  
 SQ SEQUENCE 848 AA; 94446 MW; ABFFD02ED178932F CRC64;

Query Match 3.9%; Score 80; DB 1; Length 848;  
 Best Local Similarity 25.2%; Pred No. 42;  
 Matches 58; Conservative 20; Mismatches 74; Indels 78; Gaps 15;

QY 42 ENTRNAFE-----GHVYVKGLYD-----QEGCRNDEGROVAGISLP-FDS 81  
 DB 558 FEARNDIEDFRELCKRFRNEEAFTGSLDGMGLKRMGQVQGGK-----RGVHLPAFED 613  
 QY 82 CVAARFRLNPGCIFYTTVVISF-HP-LFVTKVDRAVQCIFYEAD---KTVSAQIE 135  
 DB 614 -----FMNNKVEVEFDPQMFV-----RHQAFREDPDLEPIGTSGLIE 652  
 QY 136 VSEITTAFTQ-----DIVPVCVREYELDGGPFGOPVAFIIGOPY-----HKMCDSET 186  
 DB 653 IYSKTIADNNYDCCGHPWMEFKIERSHGSPQSQYPLHL--QSVHPDFRLHSQLESET 710  
 QY 187 VDFPCAIV--HSCFYVD-----GNGDFVEILNADGCAL-----DKY 220  
 DB 711 LMQQTVAGKEVFVFINPDASARGIRNGDVYAVFANRGVLAGAVVSDRY 760

RESULT 32  
 AND\_HUMAN  
 ID AND\_HUMAN STANDARD; PRT; 974 AA.  
 AC P19021.  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Peptidyl-glycine alpha-amidating monooxygenase precursor (EC 1.14.17.3) (PAM).  
 GN PAM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thyroid carcinoma;  
 RX MEDLINE=90290494; PubMed=2357221;  
 RA Glauder J., Ragg H., Rauch J., Engels J.W.;  
 RA "Human peptidylglycine alpha-amidating monooxygenase: cDNA, cloning and functional expression of a truncated form in COS cells,"  
 RT Biochem. Biophys. Res. Commun. 169:551-558(1990).  
 RL [2]  
 RN SUPFAM0N  
 RX MEDLINE=94193805; PubMed=8144680;  
 RA Yun H.Y., Keutmann H.T., Bipper B.A.;  
 RA "Alternative splicing governs sulfation of tyrosine or oligosaccharide on peptidylglycine alpha-amidating monooxygenase,"  
 RT J. Biol. Chem. 269:10946-10955(1994).  
 RL -1- FUNCTION: C-terminal alpha-amidation of peptides. The reaction produces a peptidyl(2-hydroxyglycine) intermediate is unstable and dismutates to glycylate and the corresponding desglycine peptide amide.  
 CC -1- CATALYTIC ACTIVITY: Peptidylglycine + ascorbate + O(2) = peptidyl(2-hydroxyglycine) + dehydroascorbate + H(2)O.  
 CC -1- COFACTOR: COPPER AND ASCORBATE.  
 CC -1- SUBCELLULAR LOCATION: Secretory granules.  
 CC -1- SIMILARITY: BELONGS TO THE COPPER TYPE II, ASCORBATE-DEPENDENT MONOOXYGENASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its



-----  
 SEND AN EMAIL TO LICENSE@STU.CU) :

DR HSSP; p12283; 1ADE.  
DR InterPro; IPR001114; Asucc\_synthase.  
DR Pfam; PF00709; Adenylsucc\_synth; 1.  
DR Prodom; PD001188; Asucc\_synthase; 1.  
DR TIGRFAMs; TIGR00184; purA; 1.  
DR PROSITE; PS01266; ADENYLOSUCCLN\_SYN\_1; FALSE\_NEG.  
DR PROSITE; PS00513; ADENYLOSUCCLN\_SYN\_2; FALSE\_NEG.  
KW Purine biosynthesis; Lysase; GTP-binding; Complete proteome.  
FT NP\_BIND 12 18 GTP (POTENTIAL).  
FT FT 138 138 BY SIMILARITY.  
FT ACT\_SITE 340 AA; 37205 MM; 5726F32397F16B07 CRC64;  
QO SEQUENCE 340 AA; 37205 MM; 5726F32397F16B07 CRC64;

Query Match	3.9%	Score 79.5	DB 1	Length 340
Best Local Similarity	24.2%	Pred. No. 14		
Matches 39	Conservative 26	Mismatches 59	Indels 37	Gaps 8
QY	205	DTVELNADGALD-KYLNNLEPYTDDMAGGEAHVYKKADRSOLFQCOISITIKEPNS	263	
Db	64	EDARLLIGAGVLYDPEVFLHEHEYSK-----YVADRTGVYKAL--IERKKK	111	
QY	264	ECVRFPCSEPFYGFVKTGAAAFKFAAAOQLRLKKRRAEPENIIDVRTDITLLEISDN	323	
Db	112	EQDDSSSYLSKKISTGTG--CGPAMNERVWRTAKLREVELDPLTDV-PLFY--N	164	
QY	324	QALFVDLHRALLQHNQGPVLLAAYONGICSPFGESFMFG	364	
Db	165	EAL-----DEGEDVEEGSOG-----FGLSLYYG	188	

RESULT 3  
YEW2\_YEA



ID YEM2\_YEAST STANDARD; PRT; 1753 AA.  
AC P32634;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Hypothetical 195.4 kDa protein in RPS26B-GLC7 intergenic region.  
GN YER132C OR SYCP-OR550.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SOURCE FROM N.A.  
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,  
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SOURCE FROM N.A.  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO S. POMBE RAL2.  
CC -1- SIMILARITY: TO YEAST MD53.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: J018916; AAC03230.1; -  
CC DR PIR: S30855; S30855.  
CC DR SGI: S0000934; P01.  
CC DR InterPro: IPR001798; Kelch.  
CC DR Pfam: PF01344; Kelch. 2.  
CC KW Hypothetical protein.  
CC SO SEQUENCE 1753 AA; 195382 MW; E4252998C6C2500B CRC64;  
  
Query Match 3.9%; Score 79.5; DB 1; Length 1753;  
Best Local Similarity 21.6%; Pred. No. 1.2e+02;  
Matches 63; Conservative 37; Mismatches 94; Indels 97; Gaps 15;  
  
QY 12 LIALSY-SIPVNGVE-----GEPEIEG-----PSTITNENTRNA-FEGHV 52  
DB 1002 LLSNLSGSDIPYEAISOEYGMNNGRDEEDGDQYGCISPNIRIFSTNNININGNE 1061  
QY 53 YVAGLDQEGCRNDEGGRVAGISLPFDSQNVARTRS-----INPRGIFV--TT 99  
DB 1062 KEEFEFSKSYINNEKSRISYSNPE---SVSTNSNNNAITELPLPLPSLIPWST 1118  
QY 100 TVVISHPLFVTKVDRAYRVQCFMEADRTVAQIEVSEITTAFOQTQIYPMPCREYILD 159  
DB 1119 ASYRAAEFFFT-----AQINGKMLLAVTLDDLIMAKI-YEI-- 1155  
QY 160 GGTGQPVQFAITGQPVYHKTKDSTFTVDFCAVHVSF-----VDDGNG----- 204  
DB 1156 -----PILEYLTELVLKTIISKEGGLSVTCALNLTPQOKVSRXCNKEGKIRKQDLS 1209  
QY 205 -----DTVEI-----LNADGALDKYLLNNLEVTDLMAQEAHVYKADRS 246  
DB 1210 SESYQDTLEIKSLANIDNGYVDSYLLRNTS-----MAQSH--YTDQS 1251

ID PYR1\_YEAST STANDARD; PRT; 2214 AA.  
AC P07259;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE URA1 protein [includes: glutamine-dependent carbamoyl-phosphate  
DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)].  
GN URA2 OR YJL130C OR J0686.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SOURCE FROM N.A.  
RA STRAIN-ATCC 28583 / FL100;  
RX MEDLINE-89378778; PubMed-2570735;  
RA Souciet J.-L., Nagy M., Je Gouar M., Lacroite F., Potier S.;  
RT "Organization of the yeast URA2 gene: Identification of a defective  
RT dihydroorotase-like domain in the multifunctional carbamoylphosphate  
RT synthetase-aspartate transcarbamylase complex.";  
RL Gene 79:59-70(1989).  
RN [2]  
RP SOURCE OF 1-510 FROM N.A.  
RX MEDLINE-87286375; PubMed-3039294;  
RA Souciet J.-L., Potier S., Hubert J.-C., Lacroite F.;  
RT "Nucleotide sequence of the pyrimidine specific carbamoyl phosphate  
RT synthetase, a part of the yeast multifunctional protein encoded by  
RT the URA2 gene.";  
RL Mol. Gen. Genet. 207:314-319(1987).  
RN [3]  
RP SOURCE OF 1268-2214 FROM N.A.  
RX MEDLINE-89255278; PubMed-2498313;  
RA Nagy M., Je Gouar M., Potier S., Souciet J.-L., Herve G.;  
RT "The primary structure of the aspartate transcarbamylase region of  
RT the URA2 gene product in Saccharomyces cerevisiae. Features involved  
RT in activity and nuclear localization.";  
RL J. Biol. Chem. 264:8366-8374(1989).  
RN [4]  
RP SEQUENCE OF 1-276 FROM N.A.  
RX STRAIN-S288C / FY1679;  
RC MEDLINE-96408771; PubMed-8948101;  
RA Katoulou C., Tzeremla M., Tavernarakis N., Alexandraki D.;  
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast  
RT chromosome X reveals 14 known genes and 13 new open reading frames  
RT including homologues of genes clustered on the right arm of  
RT chromosome XI.";  
RL Yeast 12:787-797(1996).  
RN [5]  
RP SEQUENCE OF 175-2214 FROM N.A.  
RX STRAIN-S288C / FY1679;  
RC MEDLINE-96408771; PubMed-8948101;  
RA Katoulou C., Tzeremla M., Tavernarakis N., Alexandraki D.;  
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast  
RT chromosome X reveals 14 known genes and 13 new open reading frames  
RT including homologues of genes clustered on the right arm of  
RT chromosome XI.";  
RL Yeast 12:787-797(1996).  
CC -1- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE  
CC ENZYMIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,  
CC AND ATCASE).  
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +  
CC phosphate + glutamate + carbamoyl phosphate.  
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate  
CC + N-carbamoyl-L-aspartate.  
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.  
CC -1- PATHWAY: Pyrimidine biosynthesis; second step.  
CC -1- MISCELLANEOUS: GLUTAMINE AMIDOTRANSFERASE AND CPSASE  
CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE  
CC GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).  
CC -1- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO  
CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNDER SEPARATE CONTROL.  
CC -1- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPSASES.  
CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.

RESULT 35  
PYR1\_YEAST

CC	DEFECTIVE DOMAIN.
CC	-1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; M27174; AAA68280.1; ALT_INT.
DR	EMBL; X05553; CAA29068.1; -
DR	EMBL; J04711; AAA35198.1; -
DR	EMBL; Z49405; CAA89425.1; -
DR	EMBL; X87371; CAA60825.1; -
DR	EMBL; D28139; BAA05680.1; -
DR	PIR; S05767; QZBYU2.
DR	HSSP; P00968; 1N9X.
DR	SGD; S000366; URA2.
DR	InterPro; IPR002029; Asp/Orn_Cotranf.
DR	InterPro; IPR002082; Asp_Carmlttransf.
DR	InterPro; IPR001317; CPS_GATase.
DR	InterPro; IPR000901; CPSase.
DR	InterPro; IPR002474; CPSase_sm_chain.
DR	InterPro; IPR002195; Dihydroorotase.
DR	InterPro; IPR004362; MGS_like.
DR	Pfam; PF00117; GATase; 1.
DR	Pfam; PF00165; OTCase; 1.
DR	Pfam; PF00289; CPSase_L_chain; 2.
DR	Pfam; PF00744; Dihydroorotase; 1.
DR	Pfam; PF00988; CPSase_sm_chain; 1.
DR	Pfam; PF02142; MGS; 1.
DR	Pfam; PF02729; OTCase_N; 1.
DR	Pfam; PF02766; CPSase_L_D2; 3.
DR	Pfam; PF02787; CPSase_L_D3; 1.
DR	PRINTS; PR00100; AOTCase.
DR	PRINTS; PR00098; CPSase.
DR	PRINTS; PR00099; CPSGATase.
DR	PRINTS; PR00096; GATase.
DR	TIGRfams; TIGR00670; asp_carb.tr. 1.
DR	PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR	PROSITE; PS00442; GATase_TYPE_1; 1.
DR	PROSITE; PS00866; CPSase_1; 2.
DR	PROSITE; PS00867; CPSase_2; 2.
KW	Pyrimidine biosynthesis; Ligase; Transferease; Multifunctional enzyme.
FT	DOMAIN 1 400
FT	DOMAIN 401 440
FT	DOMAIN 441 1482
FT	DOMAIN 1483 1492
FT	DOMAIN 1493 1821
FT	DOMAIN 1822 1909
FT	DOMAIN 1910 2214
FT	ACT_SITE 302 302
FT	ACT_SITE 386 386
FT	ACT_SITE 388 388
FT	CONFLICT 86 86
FT	CONFLICT 123 123
FT	CONFLICT 250 257
FT	CONFLICT 270 270
FT	CONFLICT 313 313
FT	CONFLICT 372 373
FT	CONFLICT 394 402
FT	CONFLICT 431 433
FT	CONFLICT 482 482
FT	CONFLICT 485 485
FT	CONFLICT 492 492
FT	CONFLICT 501 510
FT	CONFLICT 852 852
FT	CONFLICT 1411 1412
FT	CONFLICT 1582 1582
FT	CONFLICT 1588 1588

```

FT CONFLICT 1592 1592 V -> G (IN REF. 1 AND 3).
FT CONFLICT 1595 1595 S -> A (IN REF. 1 AND 3).
FT CONFLICT 1937 1937 A -> R (IN REF. 3).
FT CONFLICT 1937 1937 T -> I (IN REF. 3).
FT CONFLICT 1997 1997 H -> L (IN REF. 3).
FT CONFLICT 2039 2039 KILHAHE -> VRSMTDOCK (IN REF. 3).
FT CONFLICT 2158 2165 KILHAHE -> VRSMTDOCK (IN REF. 3).
SQ SEQUENCE 2214 AA; 245124 MW; 4C458304DEACD21 CR64;

Query Match 3.9%; Score 79.5; DB 1; Length 2214;
Best Local Similarity 18.2%; Pred. No. 1.7e+02;
Matches 70; Conservative 52; Mismatches 150; Indels 113; Gaps 16;

QY 47 AFEGHVYVKKGLIDDEGCGNDESGRQV-----AGISLPFSCNVAFRSLNPGIFVPTT 99
DB 959 AATYVLTMYTNADSHDSFDDHGVVVLGSGYYRIGSSVEYFDMCAVTAVTRLRANNI---K 1015
QY 100 TVVISFHFLEPTKVDRAVQCFYMEADKTVSAQIEVSEIITTAOTO-----IVPMPCR 154
DB 1016 TIMVNTND-----ETVSTYDDEADRLYFETINLERVLDIYEIENSSGVVSM----- 1062
QY 155 YEIIDGFGPTGPVOFAI-----ICQPYHKWTC 182
DB 1063 -----GGGTNNINAMTLHRENVKILGTSPIIDMSAENKRFKSRMLDQGVNDQPAKKEFTS 1117
QY 183 DSEVDTFC-----AVVHSCFVDDGNG-DVEIILNADGALDKYLLNNLEYPIIDLMAQG 235
DB 1118 MDE-AESFAEKVGVPLVRPSYVLSGAAMNTVYSKN---DLESILNQAVESRD---Y 1168
QY 236 EAHYVKKVADRKQLPYQCOISITIFKEPNSECCVRPOCSEPGQGVKWTGGAANKPAAAO LR 295
DB 1169 PVVITKYTEENK-----EIEMDAVARNGELMHVSEVENAGVHSGDGLT----- 1214
QY 296 LKKRSASEPNI-----IDVRTDINTLEISDDNOALPVDLRHALLQHNQOP 342
DB 1215 IVPQDPLAPEYVDRIYVATATIKAKLKTGTGYNIQFIKXNDKEIVIECNVA---SRSEF 1271
QY 343 VILAAYONGICMSPFGFSMFGLSI 367
DB 1272 FISKV--GVNLIELATKAIMGLPL 1294

RESULT 36
YF61_METJA STANDARD; PRT; 553 AA.
AC 058956;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-Oct-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1561.
GN MJ1561.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
CX NCBI_TaxID=2190;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Goeayne J.D.,
RA Kerejavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glouek A.,
RA Scott J.L., Geohaghen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Kleen H.-P., Eiser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
CC Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO M.JANNASCHII MJ0795 AND MJ1506.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

RT "Massive gene decay in the leprosy bacillus." ;

QY 358 GFSMENGSLIALIAAVITI 377

Db 654 GLAMGRTDVAIGADLLV 673

```

RESULT 38
ID CADH_MOUSE STANDARD; PRT; 827 AA.
AC 09100;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin) (BILL-
cadherin) (P130).
GN CDH17.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Swiss Webster; TISSUE=Intestine;
RA Angres B., Kim U., Tauber R.;
RT "Li-cadherin gene expression during intestinal development.";
RN Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.; SEQUENCE OF 26-33: 52-58: 74-81: 117-123 AND
RP 490-509. TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC STRAIN=BLB/c; TISSUE=Fetal Liver;
RA MEDLINE=20469471; Pubmed=10906147;
RA Ohnishi K., Shimizu T., Karasuyama H., Melchers F.;
RT "The identification of a nonclassical cadherin expressed during B cell
development and its interaction with surrogate light chain.";
RN J. Biol. Chem. 275:31134-31144(2000).
CC -1- FUNCTION: CADHERINS ARE CALCITONIN DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IS FOUND IN INTESTINE WITH
CC LOWER EXPRESSION IN SPLEEN, BONE MARROW, LUNG AND TESTIS. NO
CC EXPRESSION DETECTED IN LIVER, KIDNEY, HEART, BRAIN OR SKELETAL
CC MUSCLE. EXPRESSED IN PRECURSOR B-CELLS AND MYELOID CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES IN PRO- AND PRE-B-1
CC CELLS, DECREASES IN LARGE AND SMALL PRE-B-II CELLS, AND INCREASES
CC AGAIN IN IMMATURE AND MATURE B-CELLS.
CC -1- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF177669; A051125.1; -
DR EMBL: D87912; BAB03264.1; -
DR HSSP: P15116; INCI.
DR MGD: MGI:1095414; Cdh17.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 7.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA: 5
DR PROSITE: PS00232; CADHERIN_1; 3.
DR PROSITE: PS50268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal; Transposon.
FT CHAIN 1 25 CADHERIN-17.
FT DOMAIN 26 827 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 807 POTENTIAL.
FT DOMAIN 808 827 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 29 127 CADHERIN 1.
FT DOMAIN 128 243 CADHERIN 2.
FT DOMAIN 244 339 CADHERIN 3.
FT DOMAIN 340 448 CADHERIN 4.
FT DOMAIN 449 565 CADHERIN 5.
FT DOMAIN 566 666 CADHERIN 6.
FT DOMAIN 667 776 CADHERIN 7.
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 827 AA; 91645 MW; CDCECEA5A76E2B58 CRC64;

Query Match 3.9%; Score 79; DB 1; Length 827;
Best Local Similarity 20.2%; Pred. No. 50;
Matches 65; Conservative 36; Mismatches 93; Indels 128; Gaps 16;

QY 130 VSAQIEVEITTAFTQ-----QIVMPVCRREILDGPTGQ-----PVQ 168
Dy 2 VSAQLHFACLTLYLTCGCGEGKFGSLPKM---TFSIFGQPSQVIFQKTNPAVT 58
QY 169 FAITGQP-----VYHKTCDSFT-----VDFCAVAVHSCFVDGNGDTV 207
Dy 59 FELTGETDGIPIKIKDGLVYTRALDRFRAVHHLQLAALDS---HGATVDGPVPTI 113
QY 208 EILNADGALDKYLLNLEPYTDLMAQEAHV-----YRYAD-----RSOL 248
Dy 114 EVKVI-----NDNRTFLQSKYEGSVRONSRGRKFMVYNATDLDDPATPNQOL 162
QY 249 FYQCQISITTEKPNSECVPCQSPQGFGA-VKTGAAPKPAALQRLKRS----- 301
Dy 163 FYQIVIDL-----PQINDVMYFOIDSKTGAISLTPBEGSOELDPVKNPSYNIUVS 211
QY 302 -----APEN-----IIDVRTDINTLEISDNDALPDLRHRALLQNGCPVILAQV- 349
Dy 212 VKDMGQSENSESTTYVDVISIREINWKAPE-----PVEIRENSTDPH---PIKITOVW 263
QY 350 -----NGICSPFGFSM 361
Dy 264 NDPQAOYSIVKKEKLSPEPFSI 285

RESULT 39
ID TYCB_BACBR STANDARD; PRT; 3587 AA.
AC 030408;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrocidine synthetase II [includes: Arg-dependent proline adenylase
DE (Proa) (Proline activase); Arg-dependent phenylalanine adenylase
DE (Phea) (Phenylalanine activase); Arg-dependent D-phenylalanine
DE adenylase (D-Phea) (D-phenylalanine activase); Phenylalanine racemase
DE [Arg-hydrolyzing] (EC 5.1.1.11)].
GN TYCB.
OS Bacillus brevis.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OX NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 8185;
RA MEDLINE=98012987; Pubmed=9352938;
RA Mootz H.D., Marahiel M.A.;
RT "The tyrocidine biosynthesis operon of Bacillus brevis: Complete
RT nucleotide sequence and biochemical characterization of functional
RT internal adenylation domains.";
RT J. Bacteriol. 179:6843-6850(1997).
CC -1- FUNCTION: ACTIVATES THE SECOND TO FOURTH AMINO ACIDS IN TYROCIDINE
CC (IN TYROCIDINE A, PRO, PHE, AND D-PHE) AND PRIMERIZES THE LAST
CC ONE.

```



```

Db      79 LFAKMEDTDEV-----GWALDVLSPNLISTSMLGKVKYNLSLSDT-----ATGLIQYA 128
QY      244 -----DRSOLFYOQOISITIKERNSECVRPOCSEPOCFGAVKTGGAAPK-----AAAQLR 295
Db      129 LDGCVNVTQVEVD-----TVGMPEYQARLQOSEFGIEVTVKAKADALYPVSAASICAK 183
QY      296 LKKRSAEPENITDVRTDITWLEISDDNQALPYDLRHRALLQHNQOPVILAAVQNGICMS 355
Db      184 VARDQAVKKMOVEKLODLDT-----DYGSGYPNDPKTKAWLKEHVEPV----- 227
QY      356 PFGFSMEFGLS 366
Db      228 -FGFPQFVRFES 237

```

Search completed: March 24, 2003, 08:49:56  
 Job time : 34 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2003, 08:50:43 ; Search time 35 Seconds  
(without alignments)  
2278.292 Million cell updates/sec

Title: US-10-054-562a-4

Perfect score: 2026

Sequence: 1 MWIRLAFCTTLALSYRIP.....ALIAVITIRISRFRENOKA 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertibrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1653.5	81.6	385	093115	093115 ascaris lum
2	1575	77.7	390	09XYU9	09XYU9 wuchereria
3	1382.5	68.2	389	019707	019707 caenorhabdi
4	1310.5	64.7	433	018479	018479 meloidogyne
5	725.5	35.8	364	09XYM7	09XYM7 caenorhabdi
6	662	32.7	127	096775	096775 brugia mala
7	646	31.9	290	021808	021808 caenorhabdi
8	602.5	29.7	550	021540	021540 caenorhabdi
9	595	29.4	484	019053	019053 caenorhabdi
10	578.5	28.6	315	093532	093532 caenorhabdi
11	546.5	27.0	384	022164	022164 caenorhabdi
12	535	26.4	403	090250	090250 caenorhabdi
13	511	25.2	610	09XYN2	09XYN2 caenorhabdi
14	477.5	23.6	440	093967	093967 caenorhabdi
15	476	23.5	349	023097	023097 caenorhabdi
16	451	22.3	609	022680	022680 caenorhabdi

17	432.5	21.3	731	5	009586	009586 caenorhabdi
18	414	20.4	382	5	018213	018213 caenorhabdi
19	393	19.4	647	5	021573	021573 caenorhabdi
20	376.5	18.6	477	5	09T2E2	09T2E2 caenorhabdi
21	361	17.8	66	5	096776	096776 brugia paha
22	357.5	17.6	437	5	019304	019304 caenorhabdi
23	339	16.7	225	5	061816	061816 caenorhabdi
24	304.5	15.0	789	5	09BRQ4	09BRQ4 caenorhabdi
25	293	14.5	366	5	022627	022627 caenorhabdi
26	290.5	14.3	344	5	045890	045890 caenorhabdi
27	289.5	14.3	1286	5	09TXR6	09TXR6 caenorhabdi
28	244	12.0	809	5	09U3W7	09U3W7 drosophila
29	202.5	10.0	711	5	09NHV4	09NHV4 caenorhabdi
30	202.5	10.0	711	5	022724	022724 caenorhabdi
31	199	9.8	462	5	09W143	09W143 drosophila
32	174	8.6	284	5	09G0H4	09G0H4 caenorhabdi
33	166	8.2	604	5	09VZC5	09VZC5 drosophila
34	163	8.0	1549	5	09BMD4	09BMD4 drosophila
35	163	8.0	1557	5	09BMD5	09BMD5 drosophila
36	163	8.0	1638	5	09VEX3	09VEX3 drosophila
37	162.5	8.0	611	5	09VZE0	09VZE0 drosophila
38	151.5	7.5	514	5	016304	016304 caenorhabdi
39	150	7.4	758	5	09V9X1	09V9X1 drosophila
40	148.5	7.3	362	5	0960V7	0960V7 drosophila
41	142	7.0	833	5	09V9X0	09V9X0 drosophila
42	141	7.0	3507	5	023587	023587 caenorhabdi
43	140	6.9	692	5	09W5Z0	09W5Z0 drosophila
44	137	6.8	1262	5	018117	018117 caenorhabdi
45	135.5	6.7	744	5	09VAG2	09VAG2 drosophila

#### ALIGNMENTS

RESULT 1  
ID 093115 PRELIMINARY; PRT: 385 AA.

AC 093115;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE CUT-1-like cuticlin protein precursor.

CN ASCUT-1.  
OS Ascaris lumbricoides (common roundworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea;  
OC Ascarididae; Ascaris.  
OX NCBI\_TaxID=6252;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97390131; Pubmed=9249070;  
RA Timinouni M., Bazicalupo P.;  
RT "cut-1-like genes of Ascaris lumbricoides";  
RL Gene 193:81-87(1997).  
DR EMBL: U73005; AAB6646.1; -  
DR InterPro: IPR001507; Endoglin/CD105.  
DR SMART: SM00241; ZP: 1.  
KW Signal.

FT CHAIN 17 385 POTENTIAL.  
FT SIGNAL 17 385 CUT-1-LIKE CUTICLIN PROTEIN.  
SQ SEQUENCE 385 AA: 42408 MW: 1036834BID29286 CRC64;

Query Match 81.6%; Score 1653.5; DB 5; Length 385;  
Best Local Similarity 80.2%; Pred. No. 1.5e-151;  
Matches 311; Conservative 32; Mismatches 40; Indels 5; Gaps 4;

DB 2 MWIRLAFCTTLALSYRIPDPNGVEGEPEIECGFTSTTINPNRNATGEHYVYKGLYDQ 61  
1 MCRAYSF-LALFGIAAIPDPNGVEGEPEIECGFTSTTINPNRNATGEHYVYKGLYDQ 59  
62 GCRNDEGROVAGISLPFDSQCNVARTSLNPRGIFVTTVVISFHPLEFVKVBRAYVOC 121  
60 GCRNDEGROVAGISLPFDSQCNVARTSLNPRGIFVTTVVISFHPLEFVKVBRAYVOC 119

QY	122	FYMEADKVSQOIEVSETTTFMFOIYDMPACREIILDDGPTGPOVFOATIGOVYHKWT	181
QY	120	FYMEADKVSQOIEVSETTTFMFOIYDMPACREIILDDGPTGPOVFOATIGOVYHKWT	179
QY	182	CDSEVTDTFEGAVVHSCFVDGNSGJTEVLTNADGALDKYLLNLTNEPTDLMAGEAHVYK	241
Db	180	CDSEVTDTFEGAVVHSCFVDGNSGJTLITLNEEGCALDKYLLNLTNEPTDLMAGEAHVYK	239
QY	242	YADRSOLFQCOIISTITKEPNSECVRPQCSPOGFGAVKPGA-AAAPDAAAOLRLKKR	300
Db	240	YADRSOLFQCOIITTTIKERNSECPRPFCSPGQFGAVRPGGSLTAAPKORRCOJLTJKKS	299
QY	301	SAEPENIIDVDTINTLEISDNGALPVDLRRALLQ-NGOPVTLAAVONGICMSDFGF	355
Db	300	GGDYDNTLDVRTDSEALDISDRDALPMDLHRA-RRHQGVYTLSPANEGICMSDFGF	355
QY	360	SMFMGLSTALIAAVYIITISFEFRNOKA	387
Db	358	SMFMGLAVALAAAVVVSFEKLRPOOKA	385

RESULT 2			
09XYU9	PRELIMINARY;	PRT;	390 AA.
ID 09XYU9			
AC 09XYU9;			
DT 01-NOV-1999	(TREMBLrel. 12, Created)		
DT 01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT 01-MAR-2002	(TREMBLrel. 20, Last annotation update)		
DE	cuticulin-1.		
GN	CUT-1.		
OS	Wuchereria bancrofti.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;		
OC	Onchocercidae; Wuchereria.		
OX	NCBI_TaxID=6293;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Ramzy R., Helmy H., Adely M., Curtis K., Weil G.;		
RT	"Wuchereria bancrofti L3 cuticulin-1 cDNA."		
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF125580; AAD28743.2; "-"		
DR	InterPro: IPR001507; Endoglin/CD105.		
DR	SMART; SM00241; ZP; 1.		
SO	SEQUENCE	390 AA; 43026 MW; 2CA1E7E3D39FFA5E CRC64;	

Query Match	77.7%;	Score 1575;	DB 5;	Length 390;
Best Local Similarity	78.8%;	Pred. No. 6.2e-144;		
Matches 301; Conservative	29;	Mismatches 44;	Indels 8;	Gaps 4;

[illegible]

```
Db      361 FTLGLMIFVIVSVATVAITL 382
```

RESULT 3	
Q19707	
ID Q19707	PRELIMINARY;
19707	PRT; 389 AA

DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE F22B5.3 protein.  
 GN F22B5.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Chordata; Rhabditida; Rhabditidae.  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP Sims M.A.;  
 RA Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

RP	SEQUENCE FROM N.A.
RX	MEDLINE-99069613; PubMed-9851916;
RA	none.
RT	"Genome sequence of the nematode <i>C.elegans</i> : A platform for
RT	investigating biology.";
RL	Science 282:2012-2018 (1998).
RL	EMBL: Z50044; CAA90355.1;"
DR	InterPro: IPR001507; Endoglin/CD105.
DR	SMART: SM00241; ZP_1.
QO	SEQUENCE 389 AA; 42963 MW; 40BAB9C5CE031BB5 CRC64;

Query Match	68.28; Score 1382.5; DB 5; Length 389;
-------------	--

Matches 268; Conservative 42; Mismatches 52; Indels 23; Gaps 8;

Dy  
12 LIALSISIPVDNGVGEPEIECGTSTINFNTRNAFEGHYVKGLYDQECGRNDGGRQ  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 13 LVASVASIPVDNNVGEPEVECGTSTVNFNTRNAFEGHYVKGLFDQECRNDGGRQ

```
QY 72 VAGISLPEDSCNVARTRSLNPRGIEVTTTIVVISEHPLFVTKVDRAVYQCFYMEADKTVS 13
      |||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 73 VAGIELPPTCNVARTRSLNPRGVEFTTTVVVSFHPQFVTKVDRAVYQCFYMEADKTVS 13
```

QY	132	AQLEVEITTAFCQIVPMPCRYEILDDGPRGQPVGFALIGQPVYHKWTCDSFVDTFC	191
		::     ::     ::     ::     ::     ::     ::     ::     ::     ::	
Db	133	TQIEVSDLTATFCQIVPMPCICKYELINGGPRGPEPVGFATIGQOYVYHKWTCDSFVDTFC	192

[illegible]

QY 252 CQISITIKPEPNSCEVRPCQSEPOGFGAVKTCGAAPAAQAQLRLKKRSAP-PENITDV 310  
|||||:|||||||:|||||||  
b 253 CQISTTYEPNFECAEAPTCSEPOGEGAVKQANOTAO-----FFRYTKRSPDVMENITDV 307  
|||||:|||||||:|||||||

OY 311 RTDINTVEISDDNALPVDIRHALI--QHNGQPVIILAVQNGICMSPGFS--MEMG 364  
|::|||::||| |::|:  
308 PAETMTTEVTCN-IIDSGTTOLOALVASPEFGD---SEPOATCTSCPTUTWVC 361

```
QY      365 L$IALIAVIT---ISFKERNOK 386
        |::|::|:|:|
369     FM *****
```

## RESULT 4

ID	PRELIMINARY;	PRT;	433 AA.
018479			
018479;			
01-JAN-1998	(TREMblrel. 05, Created)		
01-JAN-1998	(TREMblrel. 05, Last sequence update)		
01-MAR-2002	(TREMblrel. 20, Last annotation update)		

```

DE Cuticlin-1.
GN MTCUT-1.
OS Meloidogyne artiellia.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
OX NCBI_TaxID=42426;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97218031; PubMed=9065692;
RT "Modulation of expression at the level of splicing of cut-1 RNA in the
RT infective second-stage juvenile of the plant parasitic nematode
RT Meloidogyne artiellia."
RT Mol. Gen. Genet. 253:589-598(1997).
DR EMBL: X96677; CA65452.1; -
DR InterPro: IPR001507; Endoglin/CD105.
DR SMART: SM00241; ZP: 1.
SQ SEQUENCE 433 AA; 47582 MW; 8057D216C7DCE1E8 CRC64;

Query Match 64.7%; Score 1310.5; DB 5; Length 433;
Best Local Similarity 65.8%; Pred. No. 2.9e-118;
Matches 262; Conservative 30; Mismatches 63; Indels 43; Gaps 7;

QY 2 MIRLAFCTLLALS-----YSIPVNGVGEPEIECGPTSTINFTRNAFEGHYVKG 56
DB 1 MRLLFAIGVFALNMFYTRALPVONGVGEPEIECGPTSTINFTRNAFEGHYVKG 60
QY 57 LVDEGCRNDEGRGVOAGISLPDSCNVANTRSLNPRGIFVTTTVVISHPLEVTYKVDRA 116
DB 61 LFDQACGRSDEHGRQVAGIELPFDSCNVART-DAEPGVSESTVTVISHPQETKYDRA 119
QY 117 YRVOCFEMADKTVSAQIESEITTAFOQIIVMPVRCRYELIDGPRFGQVQFAIITQPV 176
DB 120 YRVOCFEMADKTVSAQIESEITTAFOQIIVMPVRCRYELIDGPRFGQVQFAIITQPV 179
QY 177 YHWTCDSFETVDFCAVYHSCFVDNGDVEILNADGKALDKYLLNLEYPEDLMAGOE 236
DB 180 YTSGETDSFETDFCAVYHSCVVDNGDVEILNADGKALDKYLLNLEYPEDLMAGOE 239
QY 237 AHVYKADRSQLYQCOISITIKEPNSCYRPOCSFPGFAGVKTG----GAAPKPA 292
DB 240 AHVYKADRSQLYQCOISITIKEPNSCYRPOCSFPGFAGVKTG----GAAPKPA 299
QY 293 -----OLRLAKRSAP-ENLIDVRIINLEISDDN 323
DB 300 AATPPAAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 359
QY 324 QALPVDLRHRLAQHN--GQPVILAAVQNG--ICMSPF 357
DB 360 FNLPAHRIHHHLMENLSAPQRRTFYQGGQGVCMSSF 397

RESULT 5
Q9XVM7 PRELIMINARY; PRT; 364 AA.
AC Q9XVM7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F53F1.1 protein.
GN F53F1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX Burton J.;
RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]
SQ SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;

```

```

RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
DR EMBL: Z81088; CAB03124.1; -
DR InterPro: IPR001507; Endoglin/CD105.
DR SMART: SM00241; ZP: 1.
SQ SEQUENCE 364 AA; 40843 MW; D28FE2054AD97EE4 CRC64;

Query Match 35.8%; Score 725.5; DB 5; Length 364;
Best Local Similarity 47.6%; Pred. No. 1e-61;
Matches 129; Conservative 51; Mismatches 84; Indels 7; Gaps 1;

QY 19 IPVNGVEGEPEIECGPTSTINFTRNAFEGHYVGLYDOEGCRNDEGRGVOAGISLP 78
DB 49 VPIONSLYGVQVBCDSRTISVQIKTEKPEYGVAFVDPASEVCTSGRGRSAFLEIE 108
QY 79 FDSQNVARTSLNDRGIFVTTTVVISHPLEVTYKVDRAVYRQCFYMEADKTVSAQIEVSE 138
DB 109 IGLGALRQRVLNPKGLAVRTTISFHPYFITRKVDRTYNLLCYRESQYVANNISVDE 168
QY 139 ITTAFQOIYPMVRCRYEILIDGPTGQVQFAIIGQVYHKMTCDS-----TVDTFG 191
DB 169 ISTISYVNLMTPTCTYQIISGSGPGEVPEGLIGQOVYHQWKCDNDKGMITLIKEDSFC 228
QY 192 AVHSCFVDDGNGDVEILNADGKALDKYLLNLEYPEDLMAGQEAHYKADRSQLYQ 251
DB 229 MVHTCSVDGGRGRTSLIDSGCSIDKFLSLNLEFGNLLAGQEAHYKADRDALFQ 288
QY 252 QGISITIKPEPNSCYRPOCSFPGFAGVKTG 282
DB 289 QGISITIKPEPNSCYRPOCSFPGFAGVKTG 319

RESULT 6
Q96775 PRELIMINARY; PRT; 127 AA.
AC Q96775;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cuticlin (Fragment).
GN CUT-1.
OS Brugia malayi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99339397; PubMed=10413052;
RA Lewis E.K., Hunter S.J., Tellez L., Pavla Nunes C., Bazzicalupo P.,
RA Devaney E.;
RT "Cut-1-like genes are present in the filarial nematodes Brugia pahangi
RT and Brugia malayi and, as in other nematodes, code for components of
RT the cuticle.";
RL Mol. Biochem. Parasitol. 101:173-183(1999).
DR EMBL: AJ012617; CA110074.1; -
DR InterPro: IPR001507; Endoglin/CD105.
DR SMART: SM00241; ZP: 1.
FT NON_TER 1
FT TER 127
SQ SEQUENCE 127 AA; 14119 MW; 193863B7F684800 CRC64;

Query Match 32.7%; Score 662; DB 5; Length 127;
Best Local Similarity 96.1%; Pred. No. 3.6e-56;
Matches 122; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 102 VISEHPLFVTVDRAYRQCFYMEADKTVSAQIESEITTAFOQIIVMPVRCRYEILIDG 161
DB 1 VISEHPLFVTVDRAYRQCFYMEADKTVSAQIESEITTAFOQIIVMPVRCRYEILIDG 60
QY 162 PTGQPVQFAIIGQVYHKMTCDSFVDFCAVYHSCFVDNGDVEILNADGKALDKYL 221
DB 61 PTGQPVQFAIIGQVYHKMTCDSFVDFCAVYHSCFVDNGDVEILNADGKALDKYL 120

```

QY	222	LNNLEYP	228
Db	121	LNNLEYP	127

RESULT 7  
Q21808

RESULT 8		
Q21540		
TD Q21540	PRELIMINARY:	PRT: 550 AA.
AC Q21540:		
DT 01-NOV-1996	(TREMBLrel. 01,	Created)
DT 01-NOV-1996	(TREMBLrel. 01,	Last sequence update)
DT 01-MAR-2002	(TREMBLrel. 20,	Last annotation update)
DE M142.2	protein.	
EN M142.2.		

OS *Caenorhabditis elegans*.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Peltoderinae; Caenorhabditis.  
OX NCBI\_TaxId=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McMurray A.A.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode *C.elegans*: A platform for  
RT investigating biology".  
RL Science 282:2012-2016(1998).  
EMBL Z73428; CA87806.1;  
DR InterPro: IPR001507; EndoglIn/CD105.  
DR InterPro: IPR002035; WVF\_A.  
DR Pfam: PF00092; vwa; 1.  
DR PRINTS: PR00453; VWFADOMAIN.  
DR SMART: SM00327; vwa; 1.  
DR SMART: SM00241; zp; 1.  
DR PROSITE: PSS0234; vwfA; 1.  
SQ SEQUENCE 550 AA; 61445 MW; 75EDB63426BBA74A CRC64;

ID	Q19053	PRELIMINARY;	PRT;	484 AA.
AC	Q19053;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-JAN-1998 (TREMBLrel. 05, last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, last annotation update)			
DE	E04D5.3 protein.			
GN	E04D5.3.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea			
CC	Rhabditidae; Pelodolerinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	McMurray A.A.;			
RL	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RP	SEQUENCE FROM N.A.			

RX	MEDLINE-99069613; PubMed=9851916;
RA	none;
RT	"Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL	Science 282:2012-2018(1998).
DR	EMBL; Z66496; CAAG1280.1; "
DR	InterPro; IPR001507; EndoglIn/CDI05.
DR	SMART; SMO0241; ZP; 1.
SQ	SEQUENCE 484 AA; 53325 MW; DD6737B6D7E7F3B CRC64;
Query Match	29.4%; Score 595; DB 5; Length 484;
Best Local Similarity	45.0%; Pred. No. 6.4e-49;
Matches 113; Conservative	47; Mismatches 85; Indels 6; Gaps 4;
OY	27 GEPEIECGPTSTITINFNFRNAFEHGVVYKGLDQGCGRNDESGROVAGISLPFDSCNVAR 86       :    :       :       :
Dd	14 GEPEVCCETASISLLEFKTRNSFGKVFPVKGYSEPSCKTVGGKKGHREVAHDGCVAR 73     :    :  ::    :       :
OY	87 TRSLNPGEIVTTVTVVISHPLEVTWKVDNAVOCRFMEADKTVSQILEVEITTAFOTQ 146   :   ::   :     :   :            :    :   :
Dd	74 QREIN--CVMISAFTIIISFHSSIFIKRIDRAYEVSCFYEGTKKHVNHVDSALTTOLES 131   :   :    :                   :
OY	147 IVPMVCGERELIDGGPGTPGOVPQALIGQPVRHKMDCSTVDTFCAYVNSCFVDDG--NGD 205 :        :       :       :       :
Dd	132 ETQLPVCRRIILINA--GSPIKYARIIGDVHHKMTVCVALENVVCMKHSCVTYDGQGSP 190 :        :       :       :       :
OY	206 TVEILNADGCAIDLKYLNNLEYPTDLIMAGOEAHYKRVADRSQLFFOCSIITKEPNSEC 265   :     :       :       :       :       :
Dd	191 PVTVIDANGCSVDGIIONLETYSLDLTAKGLAPVKRPDKADLYRNCQIQTLTIKDYNYGC 250   :     :       :       :       :       :
OY	266 --VRFQCSEPQ 274 :    
Dd	251 SNTOPQCPETSQ 261 :    
RESULT 10	
ID	Q93532 PRELIMINARY; PRZ; 315 AA.
AC	Q93532;
DT	01-FEB-1997 (TREMBLrel_02, Created)
DT	01-FEB-1997 (TREMBLrel_02, Last sequence update)
DT	01-MAR-2002 (TREMBLrel_20, last annotation update)
DE	F2ODU1.8 protein.
GN	F2ODU1.8.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC	Rhabdilitide; Peloderineae; Caenornabditis.
OX	NCB1-TaxID=6239;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Burton J.;
RL	Submitted (Aug-1996) to the EMBL/Genebank/DBJ databases. [2]
RR	SEQUENCE FROM N.A.
RX	MEDLINE-99069613; Pubmed=9851916; none;
RA	none;
RT	"Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL	Science 282:2012-2018(1998).
DR	EMBL; Z78542; CAB01742.1; "
DR	InterPro; IPR001507; Endoglin/CDI05.
DR	SMART; SMO0241; ZP; 1.
SQ	SEQUENCE 315 AA; 35785 MW; FEDF94021B39CF00 CRC64;
Query Match	28.6%; Score 578.5; DB 5; Length 315;
Best Local Similarity	40.2%; Pred. No. 1.4e-47;
Matches 113; Conservative	47; Mismatches 76; Indels 45; Gaps 5;
OY	4 RLIACTLIATLASIPDNVGEEPEICEGCTSTINENTENAFEHGYYVGLVDQECG 63 ::     :                                     :
Dd	30 RIPFLCK-----YHAQDINGLOGEPLIRCGSESLSINFKTGAAGEGVVYGHTSMKHC 84   :     :                                     :
OY	64 RNDEGGROVAGISLPFDSCNVARTSLNRPIGIVTTTVTVVISHPLEVTWKVDNAVVOCHF 123

DB	Accession	Species	Length	Score	DB 5	Length 384	Best Local Similarity	Matches 123	Conservative 75	Mismatches 155	Indels 41	Gaps 9
Db	85	RFDATLESQVNLTVTSYISACDVIIRKRSNKGIMTATFIILSHPMFTIKIDKSYKVOCFY	144	27.0%	546.5	DB 5	384	31.2%	Pred. No. 2.3e-44			
Qy	124	MEADKTVSAQIEVS-----EITTAFOFOI-----	147									
Db	145	AEAGCTTVYQQLNVLDIAKEDEKFIIFWVGDEEGCTVSHTTGGDKKILHLKLNDSPEERISYN	204									
Qy	148	VPMVPCRYEILIDGGFTGQGPVPAITIGQPYVHKWTCDS--ETVDIFPCAVVHSCFPDDGNG	204									
Db	205	VPLPQCKRVLRESKT--EEVAFATVGQIYVHMSCEAPQNOTSEPCVYVHSCNVKDETG	263									
Qy	205	DTVELLNADGCLDKYLLNNLEXPDIAMAGCAHYKKADR	245									
Db	264	KEVQIFDENGCAVDKYLLNNLEYSDDLIGQLSYQVCSWTVR	304									
RESULT 11												
AC	Q22164	PRELIMINARY	PRT	384	AA							
ID	Q22164											
DT	01-NOV-1996	(TREMBLrel. 01, Created)										
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)										
DE	01-MAR-2002	(TREMBLrel. 20, Last annotation update)										
FE	T04F8.4	protein.										
GN	T04F8.4	Caenorhabditis elegans.										
OC	Caenorhabditis elegans.											
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;											
OC	Rhabditidae; Peloderinae; Caenorhabditis.											
OX	NCBI_TaxID=6239;											
RN	[1]											
RP	SEQUENCE FROM N.A.											
RA	Lennard N.;											
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.											
RM	[2]											
RP	SEQUENCE FROM N.A.											
RX	MEDLINE=99069613; PubMed=9851916;											
RA	none;											
RT	"Genome sequence of the nematode C. elegans: A platform for											
RT	investigating biology."											
RL	Science 282:2012-2018(1998).											
DR	EMBL; Z66565; CA91480.1; "											
DR	InterPro: IPR001507; Endoglin/CD105.											
DR	SMART: SM00241; ZP: 1.											
SD	SEQUENCE 384 AA; 43964 MW; 1B09AC47A1AA74E9 CRC64;											
Query Match												
Best Local Similarity 27.0%; Score 546.5; DB 5; Length 384;												
Matches 123; Conservative 75; Mismatches 155; Indels 41; Gaps 9;												
Qy	5	LIACFTLLALSYSPVNGVGECEPEIEGPRISITINFTNRAFEGHVYVKRLYOQEGCR	64									
Db	10	LVETFYFVLFGEFAAQQDDNELGPELIDCNADTIDMOFTRKQFGKYVKSYSNPECR	69									
Qy	65	-----NDGGROVAGISLPEFDCNVAARTSLNPRGIFVTVVVISFHPLETVKVRAYRV	119									
Db	70	VDYSKRDQGRGVGGIKIKLNHGACMNDRQMLRPEGMSPVLLISFHPLETRMDKAYRI	129									
Qy	120	OCFYEADKTVSAQIEVSEITTAFOFOIYPMVPCRYEILIDGGFTGQGPVPAITIGQPYVHK	179									
Db	130	RCMYEAKARTVYAAIDVSNLPTESVOSDLPPTCSYTTIRRDLDPIIKYKAAVQDVVHR	189									
Qy	160	WTCDEYDTFCVAVVHSCVPDGDGNDIYELIADGCLDKYLLNNLEYPDTL-MAGOEAH	238									
Db	190	WQCDSE---DGLVHSCVDEGQGEKOMTIDRSCHTDRLLLDGFTYVVALNMAVRESF										

QY 356 P-----EGFSMFMGISIALIAVITITISFKRPNO 385  
 DB 342 PKLLVAVVEFTFEFLVFTTILVAVVHYRCKNE 375

## RESULT 12

Q90250 ID Q90250 PRELIMINARY; PRT; 403 AA.

AC 090250;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Y53H1B.1 protein.  
 GN Y53H1B.1  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology."  
 RT Science 282:2012-2018(1998).  
 DR EMBL; AL132851; CAB60411.1; -  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR SMART; SM00241; 2P; 1.  
 SQ SEQUENCE 403 AA; 45346 MW; ECESES5419CA8428 CRC64;

Query Match 26.4%; Score 535; DB 5; Length 403;  
 Best Local Similarity 31.0%; Pred. No. 3.2e-43;

Matches 126; Conservative 77; Mismatches 160; Indels 44; Gaps 13;

QY 9 CTTLLALSYSI-----PYDNGVEGEPEIECGPTSTINFTNRAFGHYVGLYDDEG 62  
 DB 5 CYLLVLLITIOVMADRPPTLDNGISELPYVDCMDRYKLSKTKYORPHGRIFVYGMDKA 64  
 QY 63 CRNDEGROVAGISLPFD--SCNVAFTSLNP--RGIFVTTVVISFHPLEFVKVDRAVR 118  
 DB 65 CVADFTSQAKDVFTELENGACNMRQRMIGPEKGMESMVIYISFHSFTITKVDRAVR 124  
 QY 119 VQCFWEAKTVSAQIEVSEITTAFTQTVPMPCRYEILDGPTGQPVQAFALIGPVYH 178  
 DB 125 CTCEFYEAADKVVYTNKFDVSNLPTTDLIDTARMPLCTYSVRDSITGPIVEFAKVGETVYH 184  
 QY 179 KMTCDSEFTVDFCAVAVHSCFVDDGNGDVE--ILNADGCAIDKYLNNLEYPFD-LMAGOE 236  
 DB 185 VNNVCES---DMPSMLVHSCFVDDGNGDERKPLIDEGCAIDPLILDITLYNKDNVAYAQ 241  
 QY 237 AHVYKADSQLEFYQCQISITTIKEPNSC---VRQCSPEQFGAVTGGAAAKPPAAQA 293  
 DB 242 VMTFKFADRVSTYFCQAVS--TCMNTGCMCDGKTPRCGPAGSP-----RSSSSSN 290  
 QY 294 LRLTKRSAPENIIDVRDINTMLEISD-----NOALVVDLRHRLALHNGQPV---L 345  
 DB 291 DNGFNRRSRRA--TLNKRRAVMMVHLADTMDLAANKITVPELEEKNSDCKDEPPRAHSHL 348  
 QY 346 AAVONG-----ICMSPFGSMFGISIALIAVITISFKRPNO 387  
 DB 349 ALIQGSHPTVVCVCLNGSRIOIFVLSFLMLTVTVTLITTVKRSMTA 395

RESULT 13  
 Q9XVNZ ID Q9XVNZ PRELIMINARY; PRT; 610 AA.  
 AC 09XVNZ;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE F53B6.6 protein.  
 GN F53B6.6  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA White S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology."  
 RT Science 282:2012-2018(1998).  
 DR EMBL; Z81086; CAB03116.1; -  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR SMART; SM00241; 2P; 1.  
 SQ SEQUENCE 610 AA; 68585 MW; 09819DC5C3741329 CRC64;

Query Match 25.2%; Score 511; DB 5; Length 610;  
 Best Local Similarity 37.4%; Pred. No. 1.2e-40;

Matches 101; Conservative 58; Mismatches 101; Indels 10; Gaps 6;

QY 10 TLLINLSYS--IPYDNGVEGEPEIECGPTSTINFTNRAFGHYVGLYDDEG--RN 65  
 DB 61 TLLILSTSSCFEIONGVVGKPEVFCGIDTIVKVTNTEFGRIYVDESKOCHVOHS 120  
 QY 66 DEGGROVAGISLPDSCNVAFTSLNP-RGIFVTTVVISFHPLEFVKVDRAVRVOCFWE 125  
 DB 121 ADAHSSPQEFETIPIGACNMRQRTLHPRISISFTMTISFHFYTGMDRAISINCFLE 180  
 QY 126 ADKTVSAQIEVSEITTAFTQTVPMPCRYEILDGPTGQPVQAFALIGPVYHKTCDSE 185  
 DB 181 SIKGLNAEIDVGTLPAPQHVDOEYSLPVCAYHMKD--GIEGHVYLFPAOVGQKYNVWRCD 239  
 QY 186 TYDTFCAVAVHSCFVDDGNGDVEILNADGCAIDKYLNNLEYPFD-LMAGQAAHYKYAD 244  
 DB 240 ASHVYGIILHSCYADGHNKRELVDGSGYDPLPQIEYEHGALISAYTNAAHYKYAD 299  
 QY 245 RSQLEFYQCQISITTIKEPNSC---VRQCS 271  
 DB 300 KYQVFTCTVQCLYKH--DGCCEGITRPQCS 328

## RESULT 14

Q93967 ID Q93967 PRELIMINARY; PRT; 440 AA.

AC 093967; 094405;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE ZK265.8 protein.  
 GN ZK265.8  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dobson R.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson M., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier W., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RESULT	15			
Q23097				
ID	Q23097	PRELIMINARY;	PRT;	349 AA.
AC	Q23097;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	W01A8.3 protein.			
GN	W01A8.3.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditia; Rhabditidae; Rhabditidae; Pelodermata; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Wilkinson J.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99069613; PubMed=9851916;			
RE	note:			
RA	"genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology.";			
RL	Science 282:2012-2018(1998).			

[illegible]







RESULT 22

ID Q19304 PRELIMINARY; PRT; 437 AA.

AC Q19304; 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

DE Hypothetical 49.1 kDa protein.

GN F10E7.10.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";

RT Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Pauley A.;

RT "The sequence of C. elegans cosmid F10E7.";

RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.;

RT "Direct Submission.";

RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: U41264; AAA82426.1;

DR InterPro: IPR001507; Endoglin/CD105.

DR SMART: SM00241; ZP: 1.

KW Hypothetical protein.

SO SEQUENCE 437 AA; 49113 MW; 033C2B12BFCB748D CRC64;

Query Match 17.6%; Score 357.5; DB 5; Length 437;

Best Local Similarity 31.6%; Pred. No. 5.5e-26;

Matches 87; Conservative 47; Mismatches 122; Indels 19; Gaps 7;

QY 1 MMTRLAFTCTLTALSTIPVDNGVGEPEIEGGPTSTINTENTRNAFECHVYVKGIDYQ 60

DB 1 MLSTIYWISFFFTTWSLKYNEIYDTPVCEEDKTIITKTSMSNPSNITVDERND 60

QY 61 EGGRNDEGGROVAGISLPEDSCNVARTRSLNPRGIFVTTVVISFPLFTKVDRAVRVQ 120

DB 61 MECVS---RNQKIEVAHDKCGVANEKTEQRPNGSTRLCIFVQLHPPLVPTESDRSYCAQ 116

QY 121 CFYMEADKTVSAQIEVS-EITTAFTQTQIVP-----MPVCRYEILDGGPTGQPVQFAITG 173

DB 117 CVY--ADSHVMKDIESTLISEAPQLQSPQDAPVMPKCNYSIRKQKGPVYQASIG 174

QY 174 QPYVHKTCDDSEVDFECANVHSCFYDDGNGDTVELINDGALDKLYLNLEYPFD-LM 232

DB 175 DSFFHWSGCDGNHNG---ILVONCHVEDQGNKILLIDNGCGIDHYVMDTPVYNGEOST 231

QY 233 AGQEAHVYKYADRSOLFYOQISITIKPEPSECVR 267

DB 232 AFQETHVFKEFAQRTVTRFICQIKTKCK--GDCCR 264

RESULT 23

ID 061816 PRELIMINARY; PRT; 225 AA.

AC 061816; 01-AUG-1998 (TReMBLrel. 07, Created)

DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)

DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

DE Hypothetical 25.2 kDa protein.

GN B0511.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";

RT Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Tin-Mollam A., Sutterer C., Ozersky P.;

RT "The sequence of C. elegans cosmid B0511.";

RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.;

RT "Direct Submission.";

RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF067608; AAC17653.1;

DR InterPro: IPR001507; Endoglin/CD105.

DR SMART: SM00241; ZP: 1.

KW Hypothetical protein.

SO SEQUENCE 225 AA; 25235 MW; 29070A6340C1292C CRC64;

Query Match 16.7%; Score 339; DB 5; Length 225;

Best Local Similarity 31.2%; Pred. No. 1.4e-24;

Matches 70; Conservative 42; Mismatches 108; Indels 4; Gaps 4;

QY 49 EGHVYVKGIDYQDEGGROVAGISLPEDSCNVARTRSLNPRGIFVTTVVISFHRPL 108

DB 4 KGRVFLVGHSDQDCVSRERGRRTTITVPRDKGVETVQHGKAGYTSVNVITSHDK 63

QY 109 FVTKVDRAVRVQCFYMEADKTVSAQIEVSEITTAFTQTQIVMPVCRYEILDGGPTGQPVQ 168

DB 64 FLTKVDRAVRVQICLVPATGQVVSXALTYVPSLLKDIQVLAEGSCSEYFV-D-VTRRPAP 122

QY 169 FAITGQPVYHKTCDDSEVDFECANVHSCFYDDGNGD-IVELINDGALDKLYLNLEYP 227

DB 123 IYHVNAPLEHVTCDSTNDLFCMYHDCVINEGSKRSKTIIDSGCSLDTRRLPNRY 182

QY 228 PTD-LMAQEAHVYKYADRSOLFYOQISITIKPEPSECVRPOC 270

DB 183 ENNKLSARVMSKAFREGDVAVVEFCNVRLDLRNGTS-CPRPRC 225

RESULT 24

ID 09BK04 PRELIMINARY; PRT; 789 AA.

AC 09BK04; 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

DE Hypothetical 89.1 kDa protein.

GN Y37811a.1

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";

RT Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Ozersky P., Strowatt C.;  
RT "The sequence of *C. elegans* cosmid Y37B11A."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC091267; AAK31565.1; -  
DR InterPro: IPR001507; Endoglin/CD105.  
DR InterPro: IPR002035; VWF\_A.  
DR Pfam: PF00092; vwa; 2.  
DR PRINTS: PRO0453; VWFADOMAIN.  
DR SMART; SM00327; VWA; 2.  
DR SMART; SM00241; ZP; 1.  
DR PROSITE; PS50234; VWF\_A; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 789 AA; 89078 MW; 1E930119092A5334 CRC64;

Query Match 15.0%; Score 304.5; DB 5; Length 789;  
Best Local Similarity 26.4%; Pred. No. 1.6e-20;  
Matches 78; Conservative 58; Mismatches 117; Indels 43; Gaps 12;

OY 25 VEDEPLEEGCPTSTITNFRNAPFEGHVVYKGLYDQEGC--RDDEGRQVAGISLPDSC 82  
DB 445 VRGATVCKCTEHSMTIVFRQALQGVAMHYHDEPECIRRTDMSREIQ-MTFEGKC 503  
OY 83 NVARTSLNPRGIFVTTVIVISFPLFVTVDRAVRYOCFMEADKTVASQIEVSETTTA 142  
DB 504 GLVKTPTADGCHGNFNTVILQFHPRLITRADGLDKSCVSSA--VPRQEDRAVYKNA 561  
OY 143 FQTOIVPMVCRYEILDGPTGPVOF-ALIGQVYHKWTCDETVDFCAVYHSCFVDD 201  
DB 562 ADTQ-----CYRRLHYSP-GQCVALLDAKVGETLYHRMADSDSPREYNY--LVHDCVY-Q 611  
OY 202 GNDQVYELLNADGCALDKYLLNN-----LEYPTDLAAGEAHYKKAQDSQLEFYQOQIS 255  
DB 612 SEKHQOILLDSNGCEVDQHELETPNTSRFKDYPEDSVYFQEMSVFRRPGDGLLFHCKIS 671

OY 256 I-TIKPENSEC--VRPOCSEPOGFGAVKTGGAAPAAARLRLKKRSAPENI 307  
DB 672 LCMANDPNAPCNSIIPKC-----PKKPVLPVROKRSVSADEM 710

RESULT 25  
O22627  
ID 022627 PRELIMINARY; PRT; 366 AA.  
AC 022627;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE T21B10.6 protein.  
GN T21B10.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBL\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Baynes C.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode *C. elegans*: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z68318; CAA82696.1; -  
DR InterPro: IPR001507; Endoglin/CD105.

DR SMART; SM00241; ZP; 1.  
SQ SEQUENCE 366 AA; 40916 MW; 38D248E42F0187F5 CRC64;

Query Match 14.5%; Score 293; DB 5; Length 366;  
Best Local Similarity 24.6%; Pred. No. 7.7e-20;  
Matches 81; Conservative 53; Mismatches 121; Indels 74; Gaps 10;

OY 75 ISLPDSCNVARPSRLNPRGIFVTTVIVISFHPFVTKVDRAVRYOCFMEADKTVASQI 134  
DB 63 ILIPHECMVPRRSLNPRGIFVTTVIVISFHPFVTKVDRAVRYOCFMEADKTVASQI 122  
OY 135 EVSEITTAFOQIYPM-----PVCRYEILDGPTGPVOFALIGQVYHKWTCDETVDF 186  
DB 123 AIGS-----PKRPPSDTKGPSCTSEVLT-SPGRLAGRALGQDYYHSMQC--QN 169  
OY 187 VDFFCVAVHSCFVDDGNGDVEIILNADGCALDKYLLNNLEYPTDLAAGEAHYKKAQDS 246  
DB 170 VYESCTMIENCELVGGE-ETHNEVIDSSGCKNHEIMQLEYNHRTVGTGVKFGVSHTS 228  
OY 247 QLFYQCOISTITKEPNSECVRPOCSEPOGFGAVKTGGAAPAAARLRLKKRSAPEN 306  
DB 229 IVYFACOVRLHPOLPTGECPRKC-----DLMRKRKEDSSQFP 266  
OY 307 IIVRPTDINTLEISDDNQALPVDLHRALLQHNQGPVI-----LA 346  
DB 267 SIDVRS--QNLEISQ-----LINTSSTMRPHTEPPRIQHTCPDINHETESIEBASKA 319  
OY 347 AVQNGICSPFGSMFNGSLIALAVII 375  
DB 320 LDEBRICAD--FRSVLVASILLTVALIL 345

RESULT 26  
O45890  
ID 045890 PRELIMINARY; PRT; 344 AA.  
AC 045890;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE W06D12.1 protein.  
GN W06D12.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBL\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Basham V.M.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode *C. elegans*: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z682073; CAB04922.1; -  
DR InterPro: IPR001507; Endoglin/CD105.  
DR PRINTS; PRO1651; SECCEXPRT.  
DR SMART; SM00241; ZP; 1.  
SQ SEQUENCE 344 AA; 38563 MW; D3D50114B7B4891A CRC64;

Query Match 14.3%; Score 290.5; DB 5; Length 344;  
Best Local Similarity 27.4%; Pred. No. 1.2e-19;  
Matches 82; Conservative 45; Mismatches 87; Indels 85; Gaps 12;

OY 151 PVCRYEILDGPTGPVOFALIGQVYHKWTCDETVDFCAVYHSCFVDDGNGDVEIL 210  
DB 20 PVCRYEILMNAQCPPLSHATVGDLYHKWSCGNNKEMKCMYVHSCVVDGQGFQGLV 79  
OY 211 MADGCALDKYLLNNLEY-PTDLAAGEAHYKKAQDSQLEFYQOQISTITKE-----P 261  
DB 80 DEHGCTLDARILKELEYNKELEAGQMSVFKFADKPTVFSCMIRVENKESAMPCVIP 139



Db 704 REBITSSNAITFEKRSKOGILANDANIKPAAQREVEDICVS-----MVGILIALVITALL 757  
 QY 376 TI 377  
 Db 758 TL 759

RESULT 29  
 Q9NHV4 PRELIMINARY; PRT; 711 AA.

AC Q9NHV4; PRELIMINARY; PRT; 711 AA.  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)  
 DE Transmembrane protein RAM-5.  
 GN RAM-5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20359254; PubMed=10899109;  
 RA Yu R.Y., Nguyen C.O., Hall D.H., Chow K.L.;  
 RT "Expression of ram-5 in the structural cell is required for sensory  
 ray morphogenesis in Caenorhabditis elegans male tail."  
 RL EMBL J.19:3542-3555(2000).  
 DR EMBL; AF218866; AAF67103.1;  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR00531; TonB\_boxC.  
 DR SMART; SM00241; 2P; 1.  
 DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
 SQ SEQUENCE 711 AA; 80883 MW; 259F9A06FE27E6 CRC64;

Query Match 10.0%; Score 202.5; DB 5; Length 711;  
 Best Local Similarity 21.6%; Pred. No. 1.1e-10;  
 Matches 63; Conservative 47; Mismatches 122; Indels 59; Gaps 10;

QY 5 LIAPCTLLIALSYIPVDNG-VEGEPEIEGPGISITINENTRNAFEGHVVYVKGILYDQGC 63  
 Db 8 VIGIIFEFVLATSTDPDHAHMGVPQVTCSAKLITVSFNTNIPFGCRISVFDKLFIPAC 67  
 QY 64 RNDEGGROVAGISLPEDSCVARTSLNPGIEFTTVVISFHLFTYTKVDRAVYOCFY 123  
 Db 68 NHDYSTNIQKNATFQMDILKCANPMLKNSRLRAYVEIGFHLVMTNSDRTFLVCL- 126  
 QY 124 MEADKTVSAQIEVSEITTAFOQIIVMPV-----CRYEIDGGPTGQPVQFALIQ 174  
 Db 127 ---DNTI-----MPIVNRAQSFADCTHLVMAEWSMSSEFQ-VGD 163  
 QY 175 PVYHKWTC---DSEIVDFCAVHSCFVDGNGDTVEILNADGALDKYLLNNLEFPTDL 231  
 Db 164 AIYHEWSCKLPNPAKTQTF---LTNCNALSONGOIHLIDENGCVIDSELMGDIVY---- 216  
 QY 232 MAGOEAHVYKYADRSLF-----YOCQISITIKENSEC---VARPCS 271  
 Db 217 ---SDHVPLKYLARARIFKFLTDKRIECLTEPCNNGSPCKDRVFPKCA 263

RESULT 30  
 Q22724 PRELIMINARY; PRT; 711 AA.  
 ID Q22724;  
 AC Q22724;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)  
 DE T24C2.1 protein.  
 GN T24C2.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Baynes C.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z68120; CAA92199.2; .  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR00531; TonB\_boxC.  
 DR SMART; SM00241; 2P; 1.  
 DR PROSITE; PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
 SQ SEQUENCE 711 AA; 80882 MW; 35AB81A46FE27E6 CRC64;

Query Match 10.0%; Score 202.5; DB 5; Length 711;  
 Best Local Similarity 21.6%; Pred. No. 1.1e-10;  
 Matches 63; Conservative 47; Mismatches 122; Indels 59; Gaps 10;

QY 5 LIAPCTLLIALSYIPVDNG-VEGEPEIEGPGISITINENTRNAFEGHVVYVKGILYDQGC 63  
 Db 8 VIGIIFEFVLATSTDPDHAHMGVPQVTCSAKLITVSFNTNIPFGCRISVFDKLFIPAC 67  
 QY 64 RNDEGGROVAGISLPEDSCVARTSLNPGIEFTTVVISFHLFTYTKVDRAVYOCFY 123  
 Db 68 NHDYSTNIQKNATFQMDILKCANPMLKNSRLRAYVEIGFHLVMTNSDRTFLVCL- 126  
 QY 124 MEADKTVSAQIEVSEITTAFOQIIVMPV-----CRYEIDGGPTGQPVQFALIQ 174  
 Db 127 ---DNTI-----MPIVNRAQSFADCTHLVMAEWSMSSEFQ-VGD 163  
 QY 175 PVYHKWTC---DSEIVDFCAVHSCFVDGNGDTVEILNADGALDKYLLNNLEFPTDL 231  
 Db 164 AIYHEWSCKLPNPAKTQTF---LTNCNALSONGOIHLIDENGCVIDSELMGDIVY---- 216  
 QY 232 MAGOEAHVYKYADRSLF-----YOCQISITIKENSEC---VARPCS 271  
 Db 217 ---SDHVPLKYLARARIFKFLTDKRIECLTEPCNNGSPCKDRVFPKCA 263

RESULT 31  
 Q9W143 PRELIMINARY; PRT; 462 AA.  
 ID Q9W143;  
 AC Q9W143;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)  
 DE CG3541 protein.  
 GN CG3541.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxId=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dev I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Mostreft A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relvert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003464; AAF4734.1; -  
 DR FLYbase: FBgn0035030; CG3541.  
 DR InterPro: IPR001507; EndoglIn/CD105.  
 DR SMART: SM00241; ZP; 1.  
 SQ SEQUENCE 462 AA; 50894 MW; 05F1AA173BEA558E CRC64;

Query Match 9.8%; Score 199; DB 5; Length 462;  
 Best Local Similarity 22.2%; Pred. No. 1.3e-10;  
 Matches 94; Conservative 55; Mismatches 150; Indels 124; Gaps 20;

QY 25 VGESEIEEC-GPTSTTI-----NFTRNAFEGHYVVKLYOEGCRNDEGGROVAG--- 74  
 DB 59 VPSVRKICLSSGSMLTITIDAPNHET-GLFSGMTYPKGLSKNSCLSEY--RDHVSRLR 115  
 QY 75 ISLPDSCNVAARTSLNPRGIFVTTTVVISFHPLEFVKVDRAVYQCFYMEADKTV---- 130  
 DB 116 YLPLRSCN-TMKKETDGGIEFFNTIVLOPHLKLITDGRGTHRCAYKSKDAAMKPKK 174  
 QY 131 -----SAQIEVSEITTAFOQIYPM 150  
 DB 175 YLRKHAQRQAFRRSDRREYGRSLDKQDDDEEDVYDANAPTOEEDVT---NNEIEM 230  
 QY 151 PVC-----REIIDGGTGPVQFAIGQPVYHKTKCDSETVDTFCVAVHSCVDGN 203  
 DB 231 PEGHMKITNDEHKIADVKIGDPLTIV-----SIDKQV--YGLHVTDCIVRDEL 279  
 QY 204 G-DTVEILNADGCAIDKYLNNLEFPTDLMAGO---EAHVYKYADRSOLFYOCSISITK 259  
 DB 280 GNGEERLVGDEGCPMDNEMGOFNTODRLANVTFPAHKFPYT--TSYVYCCNVRLLCL 337  
 QY 260 EPNSECVPRQCEPQGFAGVKTGAAPAAAOQLRLKKRSABEPNI--IDVTDINT 316  
 DB 338 EDPPTQEAPOCS-----GKRPKROAAD-----SKEDGLPATIEVSGLYV 379  
 QY 317 LEISDQNALPYDLRHRALLQHNQGPVILAAVQNCISPFQSFPMG-----LSIALIA 371  
 DB 380 NENENAN-----DSDEDAVYKEK-----TTDDALCVSQRTEFAITAIAGIILMAVYA 427

RESULT 32  
 Q9GCH4 PRELIMINARY; PRT; 284 AA.  
 AC Q9GCH4  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Hypothetical protein Y55D5A.a (Y55D5A.6 protein).  
 GN Y55D5A.A OR Y55D5A.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Felodierinae; Caenorhabditis.  
 OX NCBI\_Taxid=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium."  
 RL Science 283:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (Oct-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC084196; AAG23450.1; -  
 DR EMBL: AC084196; AAK29949.2; -  
 DR InterPro: IPR001507; EndoglIn/CD105.  
 DR SMART: SM00241; ZP; 1.  
 SQ SEQUENCE 284 AA; 32361 MW; EC459689E4661283 CRC64;

Query Match 8.6%; Score 174; DB 5; Length 284;  
 Best Local Similarity 23.9%; Pred. No. 1.8e-08;  
 Matches 73; Conservative 50; Mismatches 126; Indels 56; Gaps 15;

QY 101 VVISFHPLEFVKVDRAVYQCFYMEADKTVSAQIE--VSE---ITTAFOQIYPMPCRY 155  
 DB 8 VIVAFHPLVLTSDRAFRACFEDEFKRTKEIGLEHLIGNEFQ----LPKISM 62  
 QY 156 EILDGP---TQCPQFAL-----IGQPVYHKTKCDSETVDTFCVAVHSCFVDDGND 205  
 DB 63 HILPAGESLTTTKQNFANEQKVLNVGDPIMEFKLEQEH-GIRGIOLERSASSENK 121  
 QY 206 TVEILNADGCAIDKYLNNLEFPTDLMAGQEAHV-YKYADRSOLFYOCSISITKE-PNS 263  
 DB 122 GKKIIE-NGCSIDBELISDTTYSQFSKIYATSLAFKPEEIEVIRCAVRCVKKTEHL 180  
 QY 264 ECVRPQ---CSEPQGFAGVKTGAAPAAAOQLRLKKRSABEPNIIDVTDINTLEI 319  
 DB 181 EIIINGEEDLCSANNCGFVSPSSOR-----SRQLVGKNNNNTRTDI----- 224  
 QY 320 SPDQNALPYDLRHRALLQHNQGPVILAAVQNCISPFQSFPMGLSTALIAVY---IT 376  
 DB 225 -----IYVNGRFVEKQSSPOEA-RATSTQTEFCMPDIVY--YGLSTVVICYLLITIST 275  
 QY 377 ISFKF 381  
 DB 276 VVFKF 280

RESULT 33  
 Q9VZC5 PRELIMINARY; PRT; 604 AA.  
 ID Q9VZC5  
 AC Q9VZC5.  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE CG15020 protein.  
 GN CG15020.  
 OS *Drosophila melanogaster* (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;



[illegible]





RESULT	39		
ID	Q9V9X1	PRELIMINARY;	PRT; 758 AA.
AC	Q9V9X1;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	CG12063 Protein.		
GN	CG12063.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephyridoidae; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY.		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,		

OY	31	IECGPSTJTIMENTNNAEGHYVYVGLKDQDEGRADGROVAGISLPED-----SC	82
Db	393	IDCRSGEMITKTRTSTKLPDGKUYAKGA--PKSCAVNVNN-----SLEFDLKMYNDLEC	444
OY	83	NVART---RSLNPRGIFVTYTVVISFHPLEFYTKVDRAKYOCFYMEADKTVYSAQIEV----	136
Db	445	NVROSAVGRKYN-----DIVIGHMDIIVYSSDLGLAVSCQYDITNTYVANNNDLGVY	496
OY	137	SEITFAFOQTQIY---PMFVCRYEILIDGGFTGQPVQFAIGQVYVHKMTCDSFTVDT---F	190
Db	497	GEIETLEBEELVDSPNYIMKTTANDG---SDMKRIAEVGDPLALRF-----EIVDANSPY	549
OY	191	CAVYHSCVDDGNGDTVEI---LNDGCLADKLKLLNNLEPYDILMAGQEAHYKYA-DRSQ	247
Db	550	EIVFVELVAMQGT-DSAEITTLIDANGCPTDOYIMSAMQ-----KLANNRKY	594
OY	248	LEYOCQISITTIKEPNSSECVR-----POCSEP-----QGEFVAVK----	280
Db	595	LLSQPD---AKFPSESELVQFALVTPCLPRC-EVYICONDENNGELKSLISGRKRRSVL	650
OY	281	--TGGAAPKPAAMAOLRLIKKRSAPENIIDVRYDINTLEISD	321
Db	651	NGTDGVELAIKSEROKRDVSHQADDENILLVQ---SLOJTD	689

```
RESULT 40
0960V7 ID 0960V7 PRELIMINARY; PRT; 362 AA.
AC 0960V7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE ID32760P.
GN CG12063.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacled J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY051820; AAK93244.1; -
DR FlyBase: FBgn0039851; CG12063.
SQ SEQUENCE 362 AA; 39728 MW; FC23553BA0D43ICE CRC64;

Query Match 7.3%; Score 148.5; DB 5; Length 362;
Best Local Similarity 22.5%; Pred. No. 7.5e-06;
Matches 91; Conservative 59; Mismatches 116; Indels 139; Gaps 24;

OY 44 TRNAFEGHYVVKGLYDOEGCRNDGCGROYAGISLFPD-----SCNART---RSLNP 92
DB 7 TSKLFDGKAYAKA--PKSCAVNNN-----SLEFDLKMRYNDLECNVROSAYGRYM- 57

OY 93 RGIFVTTTVISFPLEFVTVDRAYRYQCFYMEADKTVSAQIEY---SEITTAFOQIY- 148
DB 58 -----DIYIOHMDMIVTSSDGLAVSCQIDLTNKTVMNNVDLGVTGETESTLSEIIV 110

OY 149 --PMYVCRYEILDGGPTGQVPFAIIGQPYVHKWTCDESETVDT---FCAYVHSCFVDDGN 203
DB 111 DSPNVIMKKTARDG---SDMKRIAEVGDPLALRF---ETVDANSPYEIFVRELYAMDGT 163

OY 204 GDTVEI--LNADGCALDKYLLNLEYPPTDLMAQGEAHVRYA--DRSQLYQCQISITIKE 260
DB 164 -DSEITLIDANGCPTDQYIMSAMQ-----KLANNRKVLLSQFD---AFKF 205

OY 261 PNSECVR-----PQCEP-----QGFCAVK-----TGAAPAA 291
DB 206 PSEELVOFRALVPCIPRC-EPVICNDENGELKSLISYGRKRKSVLNGTGVELAIKSE 264

OY 292 AQRLTKKRSAPENIIDVPTDINTLEISD-----DNOA-LPVDL 330
DB 265 RQKRDSHQAGADENITLVQ---SIQITKFAFNGADAPCGSGSEAGLDGLAKLOLDL 320

OY 331 RHRALLQHNQPVILAAVONGICMSPFGF---SMENGISIALIA 371
DB 321 -----GTSKSDTCINGYGFITAGALFLLQLTVIA 349
```

Search completed: March 24, 2003, 08:51:46  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**